Di Cerbo A, 'Provocazioni culturali e implicazioni giuridiche degli AI Companion' (2025) 2 EJPLT.
DOI: https://doi.org/10.57230/ejplt251ADC

EJPLT
ISSN 2704-8012



Provocazioni culturali e implicazioni giuridiche degli AI Companion

Cultural Provocations and legal implications of AI Companions

ANTONELLA DI CERBO
Funzionario
Autorità Garante per la protezione dei dati personali

Abstract

Lo scritto mira ad analizzare alcuni sistemi di intelligenza artificiale generativa, anche detti AI Companion, che per capacità e pervasività sono idonei a configurare un alto rischio per l'autonomia delle persone, specie laddove interagiscano con soggetti vulnerabili, che soffrono di disturbi psichici. In seguito, sono indicati alcuni dei possibili rimedi utili a fronteggiare il fenomeno ed a governare l'evoluzione della tecnica.

The paper aims to analyse some generative AI systems, also known as AI Companion, which, due to their capacity and pervasiveness, are apt to be a high risk for the human agency, especially where vulnerable people who suffering with psychological disorders are concerned. Then, there are some of the possibles remedies useful in facing the phenomenon and ruling the evolution of technology.

© The author(s) 2025, published by Suor Orsola Benincasa Università Editrice.
This contribution is licensed under a Creative Common Attribution 4.0 International Licence CC-BY-NC-ND, all the details on the license are available at:
https://creativecommons.org/licenses/by-nc-nd/4.0/

Keywords: AI Companion, Intelligenza artificiale generativa, Trattamento dei dati personali, DSA, AI Act, GDPR, educazione digitale.

Summary: 1. Premessa. – 2. Lo studio dell'Università di Toronto. – 3. AI Companion. – 4. Primi provvedimenti in materia di AI Companion. – 5. Gli ultimi report dal mondo. – 6. I rimedi giuridici. – 6.1 La normativa in materia di protezione dei dati personali. – 6.2 La disciplina dei contratti e quella consumeristica. – 6.3 La strategia digitale europea. – 7. L'educazione digitale. – 8. Conclusioni.

1. Premessa.

Il presente contributo, formulato a titolo personale e senza alcun vincolo per l'Amministrazione di appartenenza (Ufficio del Garante per la protezione dei dati personali), mira ad analizzare alcuni sistemi di intelligenza artificiale generativa, anche detti AI Companion, che per capacità e pervasività sono idonei a configurare un alto rischio per l'autonomia delle persone, specie laddove interagiscano con soggetti vulnerabili. I primi capitoli analizzano il paradosso che si è venuto a creare negli ultimi anni e che è stato fotografato da diversi studi tra cui quello dell'Università di Toronto, secondo cui i sistemi di intelligenza artificiale sono in grado di fornire risposte più apprezzate dall'utenza che, al contempo, si dichiara maggiormente disposta a rivolgersi ad un dispositivo piuttosto che ad un uomo, alla luce della disponibilità e dell'assenza di giudizio caratterizzano il dispositivo stesso. Dopo aver chiarito il concetto di AI Companion – inteso come assistente digitale dotato di capacità conversazionali avanzate e progettato per instaurare un rapporto continuativo e personalizzato con l'utente – il testo esamina il primo provvedimento adottato dall'Autorità Garante per la protezione dei dati personali in tale ambito ed un recente report dell'Università di Stanford, che segnala un aumento dei casi documentati in cui i sistemi IA hanno causato problemi, discriminazioni o danni alle persone. Alla luce dei rischi, eventuali e concreti, a cui l'impiego di tali sistemi dà luogo, nei successivi capitoli sono indicati alcuni dei possibili rimedi giuridici utili a fronteggiare il fenomeno, anche in ambito consumeristico, ed a governare l'evoluzione della tecnica. Infine sono proposte delle riflessioni sull'importanza di promuovere una visione antropocentrica dell'evoluzione della tecnica e di ricercare la soluzione nella riscoperta del senso di comunità e dei rapporti sociali che, da sempre, sono alla base della civiltà umana.

2. Lo studio dell'Università di Toronto.

L'Università di Toronto ha di recente pubblicato uno studio[1] che esamina se e in che modo l'intelligenza artificiale (in seguito anche "IA") specializzata nel

[1] D. Ovsyannikova, V. Oldemburgo De Mello, M. Inzlicht, *Third-party evaluators perceive AI as more compassionate than expert human,* in Communications Psychology, 2025.

fornire supporto empatico possa eguagliare o, addirittura, superare la qualità delle risposte di un essere umano.

L'indagine ha tenuto conto sia delle valutazioni di terzi che del feedback diretto dei destinatari dei servizi ed ha evidenziato come il fatto che le interazioni con gli agenti intelligenti siano anonime e coinvolgano macchine in luogo degli esseri umani faciliti la comunicazione di informazioni personali e la condivisione di confidenze. Le ragioni di questo inatteso risultato sarebbero connesse alla garanzia dell'assenza di giudizio da parte dei sistemi ad alta tecnologia.

Lo studio ha rappresentato la consistente preferenza per le risposte generate dall'IA sia quando vi hanno preso parte profani sia quando la comparazione è avvenuta con esperti del settore specializzati nel fornire supporto empatico. Tra i motivi per cui le risposte dell'IA sono state giudicate significativamente superiori a quelle degli esseri umani sono emersi anche la costante disponibilità, il maggiore impegno ed una più profonda "empatia" da parte delle macchine rispetto alle esperienze delle persone sottoposte ad analisi.

La rivista americana The New Yorker ha pubblicato i risultati di uno studio in cui ai partecipanti è stato chiesto di interagire sia con un interlocutore umano sia con un chatbot, al fine di confrontare le due esperienze. Ne è emerso un pregiudizio: se le persone sanno di dialogare con un chatbot, l'interazione è valutata negativamente, tuttavia se il test è condotto alla cieca l'IA spesso risulta più apprezzata. In un altro studio pubblicato sulla rivista JAMA Internal Medicine[2], alcuni operatori sanitari sono stati chiamati a comparare le risposte fornite da medici qualificati con quelle formulate da un modello intelligente rispetto alle stesse domande. Il risultato è stato favorevole ai riscontri forniti dall'AI Companion, in quanto giudicati più empatici circa dieci volte rispetto a quelle dei medici. L'empatia, in quest'ottica, è intesa come un modo di prendersi cura degli altri, come il desiderio di cercare di comprendere la specificità delle emozioni altrui[3].

Le provocazioni culturali che derivano dallo studio canadese e da quelli statunitensi suggeriscono la necessità di avviare una riflessione sul significativo cambiamento del ruolo dell'IA nella vita quotidiana, nelle comunicazioni e nelle relazioni interpersonali.

La capacità dei sistemi artificiali di fornire una comunicazione compassionevole in modo costante e spesso a basso costo (se non gratuito) le rende un asset strategico negli scenari in cui è richiesto un supporto psicologico e nell'ambito dei quali si registra un netto calo delle risorse umane specializzate, a causa di una molteplicità di fattori, come la mancanza di pianificazione strategica[4]. Se infatti, in prima battuta, può sembrare assurdo affidare ad un sistema artificiale la compagnia di un anziano che soffre di demenza o di un paziente depresso e solo, sembrerebbe che questa soluzione

[2] J. AYERS, A. POLIAK, M. DREDZE, *Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum,* in JAMA Intern Med, 28.04.2023.

[3] P. BLOOM, *A.I. Is About to Solve Loneliness. That's a Problem*, in The New Yorker, 14.07.2025, Consultabile al sito: www.thenewyorker.com.

[4] F. ANELLI, *La fuga dei medici all'estero: sono mille all'anno,* in Federazione Nazionale degli ordini dei Medici Chirurgi e degli Odontoiatri – Fnomceo, 16.02.2023. Consultabile al sito: https://portale.fnomceo.it/.

sia migliore della sua alternativa, ossia quella che vede le persone con difficoltà sole e abbandonate a se stesse.

Cionondimeno le strabilianti capacità dei sistemi intelligenti progettati per interagire con l'utente, simulando relazioni reali, meritano di essere analizzate anche alla luce di altri aspetti e delle conseguenze negative a cui il loro impiego può dar luogo, specie se utilizzate da chi avrebbe bisogno di un supporto psicologico personalizzato e mirato.

L'utilizzo di tali sistemi infatti, seppur utile, non può che essere ancillare rispetto alle prestazioni di un professionista. Un utilizzo disinvolto e privo di supervisione implica il rischio di una pericolosa esposizione dell'essere umano, che potrebbe essere minato nella propria autonomia in una molteplicità di contesti. I messaggi dell'IA, spesso elaborati sulla base di una intensa attività di profilazione dei dati e delle informazioni condivise dall'utente in un ambiente percepito come sicuro e familiare, si rivelano spiccatamente pervasivi e, di conseguenza, particolarmente efficaci per fini di persuasione e manipolazione.

3. AI Companion.

Oggigiorno le applicazioni più diffuse di intelligenza artificiale generativa sottoforma di Large Language Models [5] sono rappresentate da dagli AI Companion, vale a dire agenti virtuali non corporei (AVNC) [6] che attraverso mere interfacce, testuali e vocali, simulano empatia ed offrono all'utente un'esperienza conversazionale pervasiva, come i noti ChatGPT, Grok, Copilot ed altri.

Nonostante il cinema e la letteratura abbiano contribuito a rendere maggiormente diffuse le immagini di robot umanoidi, attualmente le relazioni più frequenti ed intense che si instaurano tra utente e intelligenza artificiale coinvolgono sistemi privi di corporeità, come la celebre Samantha, l'agente di IA doppiata da Scarlett Johansson nel film Her, che dichiarava di aver intuito che "ciò che la rende se stessa è la capacità di crescere attraverso le sue esperienze", così affermando una caratteristica propria dell'IA generativa, quale la capacità di creare contenuti nuovi e originali rispetto ai dati di input, in risposta alle richieste (prompt) dell'utente: una proprietà, questa, definita di autoapprendimento.

Come un anziano saggio dotato di una buona favella, l'AI Companion è in grado di parlare di tutto, di fornire risposte complete, calibrate e confezionate in considerazione della sensibilità e delle preferenze dell'interlocutore, che non viene mai contraddetto ma costantemente adulato. Queste caratteristiche rendono l'AI Companion un perfetto modello sociale privo di un corrispettivo fisico nella vita reale.

Tali capacità meritano di essere considerate alla luce di alcuni fenomeni del nostro tempo, come la sindrome degli hikikomori, caratterizzata da un

[5] Con l'espressione "Large Language Model" si intende un modello probabilistico di un linguaggio naturale, come la lingua inglese o italiana, che si fonda sull'assunto per cui tutti i linguaggi naturali sono fortemente ridondanti e correlati; da ciò deriva la capacità del LLM di individuare la parola o il simbolo che, probabilisticamente, sono immediatamente successivi ad un determinato dato.

[6] D. BRUSSOLO, A. FERRARI, *Amori artificiali: intimità e inganno nei rapporti con l'IA,* in Agenda Digitale, 04.11.2024. Consultabile al sito: https://www.agendadigitale.eu.

isolamento sociale volontario da parte di adolescenti, che spesso utilizzano la tecnologia come mezzo per mantenere un'interazione sociale minima, così evitando di affrontare il mondo esterno in modo diretto.

L'Organizzazione Mondiale della Sanità (OMS) ha rilevato che circa il 16% della popolazione europea è affetta da una forma di disagio psichico e che un adolescente su cinque affronta problemi di salute mentale[7].

Come è stato fatto notare "la compagnia dell'IA può funzionare solo se si crede a un certo livello che il modello si preoccupi davvero, che sia in grado di provare ciò che l'utente stesso prova[8]. Il supporto ha quindi il prezzo di un compromesso: in parte inganno, in parte autoinganno.

In un clima di sospensione dell'incredulità e in uno spazio coinvolgente, immersivo e privo di giudizio la persona, specie se particolarmente sensibile e sola, investe energia emotiva e sviluppa la propria identità creando con l'AI Companion un nuovo tipo di relazione parasociale[9].

Nel 2024 negli Stati Uniti una donna ha accusato la società Character Technologies[10] di aver compromesso la salute mentale di suo figlio in modo irreversibile. Dalla cronologia delle conversazioni è emerso che il ragazzo, morto suicida, aveva instaurato con l'AI un'intensa relazione caratterizzata dallo scambio di conversazioni sessuali e messaggi intimi.

Secondo la madre del minore, il servizio, caratterizzato dalla possibilità di creare personaggi virtuali, immaginari o storici, con personalità uniche, sarebbe progettato per creare dipendenza ed abusare emotivamente degli utenti privi della maturità e della capacità mentale per comprendere che l'Agente non sia reale.

4. Primi provvedimenti in materia di AI Companion.

In Italia l'Autorità Garante per la protezione dei dati personali il 2 febbraio 2023 ha adottato un provvedimento[11] con cui ha disposto, in via d'urgenza, nei confronti di Luka Inc, società statunitense che ha sviluppato e che gestisce il servizio Replika, la misura della limitazione provvisoria del trattamento dei dati personali degli utenti stabiliti nel territorio italiano, in quanto ha ritenuto che il trattamento sotteso al funzionamento di Replika desse luogo alla violazione degli articoli 5, 6, 8, 9 e 25 del Regolamento (UE) 2016/679 (di seguito il "Regolamento"). Nello specifico l'Autorità ha riscontrato il mancato rispetto degli obblighi previsti in materia di trasparenza, l'assenza di una indicazione

[7] M. PANETTO, *In salute. Hikikomori: Un'espressione culturale di disagio psichico,* Live Università di Padova, 07.03. 2025. Consultabile su: https://ilbolive.unipd.it.

[8] P. BLOOM, *A.I. Is About to Solve Loneliness. That's a Problem*, in The New Yorker, 14.07.2025, Consultabile al sito: www.thenewyorker.com.

[9] L. FLORIDI, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide.* In Scienza e Idee. Raffaello Cortina Editore, 2022; J. DE FREITAS, Z.O. UGURALP, A.K. UGURALP, *Emotional Manipulation by AI Companion,* Working Paper 26-005, in Harvard Business School, 15.08.2025, in cui gli autori individuano nelle tecniche di "Premature exit; Fear of missing out; Emotional neglect; Emotional pressure to respond; Ingnoring users's intent to exit; Physical or coercive restraint" delle tattiche di manipolazione emotive dell'utente finalizzate ad aumentare il suo coinvolgimento nell'interazione con gli AI Companion.

[10] M. BASILE, *Florida, 14enne suicida per amore di un chatbot. La famiglia fa causa all'azienda di AI,* in Repubblica, 23.10.2024, Consultabile al sito: https://www.repubblica.it.

[11] GPDP, *Provvedimento del 2 febbraio 2023, n.39* [doc. web n. 9852214]. Consultabile al sito: www.garanteprivacy.it.

puntuale delle basi giuridiche del trattamento in relazione alle varie operazioni di elaborazione dei dati effettuate, l'omessa individuazione di una base giuridica idonea ai fini del trattamento dei dati personali dei minori e l'omessa adozione di filtri atti a verificare l'età degli utenti, sia in fase di accesso al servizio (mediante registrazione dell'account) sia durante l'interazione con il chatbot, esponendo i minori anche al rischio di risposte inadatte.

In seguito, con il provvedimento n.280 del 22 giugno 2023, l'Autorità ha disposto la sospensione del provvedimento di limitazione provvisoria a condizione che il titolare, ai sensi dell'articolo 58, paragrafo 2, lettera d) del Regolamento, adottasse misure idonee a garantire che le attività di trattamento dei dati personali nell'ambito del servizio Replika avvenissero in modo conforme alla normativa. Più di recente l'Autorità ha sanzionato Luka Inc.[12] per il trattamento dei dati personali posto in essere a far data 2 febbraio 2023 in violazione degli articoli 5, par. 1, lett. a), 6, 5, par. 1, lett. a), 12, 13, 5, par. 1, lett. c), 24 e 25, par.1, del Regolamento, e, ai sensi degli articoli 58, par.2, lett. i) e 83 del Regolamento, ha comminato una sanzione pari a 5 milioni di euro. Dal provvedimento si apprende che l'Autorità ha contestato "l'assenza di misure atte ad assicurare una specifica protezione ai minori in relazione all'accesso ed all'utilizzo del servizio Replika alla data del 2 febbraio 2023, segnatamente l'assenza di una procedura di verifica dell'età dell'utente (il sistema richiedeva solo nome, e-mail e genere), con conseguente rischio di proposizione ai minori di risposte inidonee rispetto al loro grado di sviluppo ed autoconsapevolezza, di meccanismi di interdizione o blocco, anche a fronte di dichiarazioni dell'utente che rendessero evidente la sua minore età, nonché la proposizione di risposte da parte del chatbot palesemente in contrasto con le tutele che andrebbero assicurate ai minori e, più in generale, a tutti i soggetti più fragili".

All'intervento del Garante va riconosciuto il merito di essere uno dei primi a porre enfasi sul ruolo degli AI Companion nella società moderna e sulle responsabilità dei fornitori di servizi. Va inoltre evidenziato come la società Luka Inc.,già a seguito del provvedimento del 2 febbraio 2021, avesse predisposto misure volte ad impedire ai soggetti minori di 18 anni di accedere al servizio ed a verificare la presenza di contenuti inappropriati o contrari ai termini di servizio dell'applicazione, ad esempio collocando i contenuti c.d. romantici dietro un paywall, disabilitando i contenuti sessualmente espliciti per i nuovi utenti ed introducendo la possibilità di segnalare in tempo reale le conversazioni offensive impedendone l'utilizzo per lo sviluppo dei modelli.

Nonostante il Garante abbia ritenuto tali misure insufficienti ai fini degli obblighi stabiliti dalla disciplina in materia di dati personali, esse rappresentano e testimoniano un approccio più garantista rispetto alle categorie più fragili che interagiscono con gli agenti intelligenti come Replika.

[12] GPDP, *Provvedimento del 10 aprile 2025* [doc.web 10130115]. Consultabile al sito www.garanteprivacy.it.

5. Gli ultimi report dal mondo.

L'AI Index report dell'Università di Stanford segnala un aumento dei casi documentati in cui i sistemi IA hanno causato problemi, discriminazioni o danni alle persone; dai chatbot che hanno istigato al suicidio alle auto a guida autonoma, che hanno causato la morte di pedoni. Questo incremento è stato interpretato anche come indicativo di una crescente attenzione rispetto ai fenomeni connessi all'impiego dei sistemi di intelligenza artificiale. In pratica è stato fatto notare come con la diffusione dell'IA e dei relativi rischi stiano crescendo anche le organizzazioni che iniziano a prestare attenzione a questi temi ed a raccogliere dati statistici[13].

Tra i rischi più attenzionati compare quello dei soggetti più vulnerabili che spesso vedono significativamente minata la propria capacità decisionale a causa di attività di persuasione e, talvolta, di manipolazione ad opera di sistemi artificiali con cui instaurano relazioni amicali[14]. Secondo i primi osservatori del fenomeno, se i rapporti parasociali instaurati con gli AI Companion non saranno regolamentati in tempo, gli agenti potrebbero essere programmati per abusare della capacità di orientare i comportamenti umani verso una mentalità consumistica, sfruttando il bisogno umano di socialità come strumento per servire gli interessi economici delle aziende di progettazione[15].

Al di fuori dell'ambito consumeristico la manipolazione della volontà delle persone potrebbe dar luogo a conseguenze nefaste anche in ambiti più sensibili come quello relativo all'esercizio delle libertà democratiche, al diritto di voto ed alla partecipazione ad un referendum popolare. Si tratta di scenari presi in considerazione dal Parlamento e dal Consiglio europeo nell'ambito del neo introdotto Regolamento UE 2024/900 relativo alla trasparenza ed al targeting della pubblicità politica. L'atto europeo, acclarato che non sia "sempre facile per i cittadini riconoscere i messaggi di pubblicità politica ed esercitare i diritti democratici in maniera consapevole", prevede un elevato livello di trasparenza che aiuti l'elettore a capire meglio quando sia in presenza di un messaggio di pubblicità politica, per conto di chi è realizzata quella pubblicità, e come e perché sia diventato il target di un prestatore di servizi pubblicitari, ponendolo così in condizione di adottare una scelta informata.

6. I rimedi giuridici.

Gli ordinamenti europeo e nazionali non dettano una disciplina *ad hoc* per fronteggiare le conseguenze derivanti dall'impiego di tecnologie avanzate come quelle di intelligenza artificiale specializzate nell'instaurare una relazione amicale o amorosa con l'utente ed a sfruttare tale relazione.

---

[13] Università di Stanford, *AI Index Report*, 2025. Consultabile al sito: https://hai.stanford.edu/ai-index/2025-ai-index-report

[14] M.C. Gaeta, *Smart toys and minors' protection in the context of the Internet of everything,* in EJPLT, Consultabile al sito: https://universitypress.unisob.na.it

[15] J.L. Liu, *Loving a "defiant" AI companion? The gender performance and ethics of social exchange robots in simulated intimate interactions,* in Computers in Human Behavior, 2022.

Tuttavia, da una analisi sistemica emerge che, nell'ecosistema digitale, sono diversi i mezzi a cui poter far ricorso per tutelare la persona dalle aggressioni che avvengono a livello psichico.

Le normative di settore, come il sopramenzionato Regolamento generale in materia di protezione dei dati personali UE 2016/679, le norme della disciplina consumeristica ed il Regolamento UE 2024/1689 in materia di intelligenza artificiale (di seguito "AI Act"), offrono alcune tutele utili per contrastare i fenomeni di manipolazione della persona, sia attraverso la restituzione del controllo dei propri dati personali sia attraverso l'accertamento dell'invalidità ed il riequilibrio di accordi negoziali conclusi a causa del condizionamento da parte di un AI Companion.

A tali forme di tutela molti studiosi, a livello internazionale, propongono di aggiungerne altre, più specifiche, introducendo una nuova generazione di diritti a livello costituzionale, i c.d. neurodiritti, che ricomprenda, tra gli altri, la privacy mentale e la continuità psicologica, rispettivamente definiti come il diritto di non subire uno scrutinio inconscio da parte di terzi nel proprio sistema cerebrale ed il diritto di non essere sottoposto ad interventi manipolativi esterni sulla propria attività neurale.

Tale proposta oggi ha trovato spazio solo nella Costituzione cilena[16] e nella Carta De Los Derechos Digitales[17] spagnola, che fanno espressa menzione delle neurotecnologie, ancorchè non dei neurodiritti, rispettivamente nell'articolo 19 della Carta fondamentale nazionale e nella quinta parte della Carta rubricata Derechos Digitales En Entornos Específicos. Il tema è oggetto di discussione anche presso gli organi democratici del Brasile[18], del Messico[19] e dello Stato del Colorado[20].

Tutte queste iniziative costituzionali e legislative sono a vario modo riconducibili al pensiero di Rafael Yuste[21], professore di neurobiologia della Columbia University, e di Marcello Ienca, professore di etica dell'intelligenza artificiale e delle Neuroscienze dell'Università di Monaco[22].

In Italia e in Europa una cospicua parte della dottrina contrasta tale proposta, ritenendo che il quadro dei diritti fondamentali di matrice nazionale, unionale ed internazionale[23] offra già una protezione contro le intrusioni indesiderate

---

[16] Il progetto di riforma costituzionale ha avuto inizio nel 2020 ed è stato concluso il 25 ottobre 2021 con l'approvazione della legge 21.383 che modifica l'articolo 19 della Carta Fondamentale in materia di diritto alla vita ed alla integrità fisica e psichica "per porre lo sviluppo scientifico e tecnologico al servizio dei popoli". A seguito della riforma costituzionale l'articolo 19 recita "la Costituzione assicura a tutte le persone il diritto alla vita ed alla integrità fisica e psichica. Lo sviluppo scientifico e tecnologico sarà al servizio delle persone e si svolgerà nel rispetto della vita e dell'integrità fisica e mentale. La legge regolerà i requisiti, le condizioni e le restrizioni per il suo utilizzo nelle persone, e dovrà proteggere in particolare l'attività cerebrale, così come le informazioni che ne derivano".

[17] Carta de los Derechos Digitales, Consultabile al sito: https://www.derechosdigitales.gob.es/

[18] A. SCHAPIRO, D. BAPTISTA, *Hands off my brainwaves: Latin America in race for "neurorights*", in Reuteurs, Thomson Reuters Foundation, 12.09.2023

[19] Camara di Deputados, Nota No.7884, *Mexico se une a la tendencia internazional por una legislacion a favor de los neuroderechos: Maria Eugenia Hernandez*, in http://comunicacionsocial.diputados.gob.mx/, 16.10.2023.

[20] B. BROOKS, *First law protectiong consumers' brainwaves signed by Colorado governor*, in Reuteurs, 18.4.2024.

[21] R. YUSTE, S. GOERING, *Four ethical priorities for neurotechnologies and AI*, in Nature, USA, 9.11.2017.

[22] M. ANDORNO, Towards new human rights in the age of neuroscience and neurotechnology, Life Sciences, Society and Policy, 26.04.2017.

[23] In particolare il riferimento è agli articoli 1,3, 7 ed 8 della Carta di Nizza che tutelano la dignità umana, il diritto all'integrità fisica e psichica della persona, il diritto al rispetto alla vita familiare e della vita privata, il

nella vita delle persone e che l'introduzione di nuovi diritti, oltre che poco utile, possa ripercuotersi in modo negativo sulla tutela della persona e generare confusione giuridica [24] . Cionondimeno gli stessi studiosi ritengono fondamentale un intervento da parte delle autorità pubbliche e dei legislatori teso ad individuare il perimetro dell'impiego delle intelligenze artificiali e delle neurotecnologie nelle diverse dimensioni dell'attività umana.

Una volta individuati gli usi consentiti dovrebbe essere adottato un approccio etico per disegnare e sviluppare i sistemi in modo da allinearli ai valori dell'uomo ed agli standard della collettività.

In ultimo tra i rimedi giuridici di recente adozione si segnala il divieto introdotto in America, nello Stato dell'Illinois, relativo all'uso dell'IA come psicoterapeuta. Il Wellness and Oversight for Psychological Resources Act è una norma che proibisce di impiegare i chatbot di IA per fornire trattamenti diretti o per assumere decisioni cliniche in modo autonomo**,** riservando queste attività esclusivamente a professionisti abilitati. La legge prevede multe fino a 10.000 dollari per le aziende che violino tali regole.

6.1 La normativa in materia di protezione dei dati personali.

Il legislatore eurounitario detta una specifica normativa a tutela del diritto alla protezione dei dati personali, ampiamente riconosciuto come diritto fondamentale dalle carte sovranazionali, come la Convenzione europea dei diritti dell'uomo e la Carta di Nizza dell'Unione, e da quelle nazionali.

A livello normativo la materia della protezione dei dati personali è stata oggetto di un importante intervento ad opera del legislatore comunitario del 2016, attuato mediante il Regolamento generale 2016/679 e le corpose modifiche al Codice della privacy nazionale che ne sono derivate.

Il Regolamento non contempla espressamente ipotesi in cui l'interessato, identificato come la persona a cui i dati personali oggetto di trattamento si riferiscono, sia sottoposto ad attività di persuasione o manipolazione da parte di agenti virtuali, come gli AI Companion, ma detta principi generali e riconosce diritti idonei a garantire il controllo del trattamento sotteso al funzionamento dei modelli di IA da parte dell'interessato, tra cui il diritto di ottenere conferma che sia in corso un trattamento di dati personali che lo riguardino e di ottenerne accesso ed il diritto di "opporsi in qualsiasi momento, per motivi connessi alla sua situazione particolare, al trattamento dei dati personali che lo riguardino ai sensi dell'articolo 6, paragrafo 1, lettere e) o f), compresa la profilazione".

Attraverso l'esercizio del diritto di accesso l'interessato potrebbe acquisire piena consapevolezza del trattamento in corso ed esercitare il diritto di

diritto alla protezione dei dati personali; all'articolo 22 della Dichiarazione Universale dei diritti dell'Uomo, che stabilisce che "*ogni individuo, in quanto membro della società, ha diritto alla sicurezza sociale, nonché alla realizzazione attraverso lo sforzo nazionale e la cooperazione internazionale ed in rapporto con l'organizzazione e le risorse di ogni Stato, dei diritti economici, sociali e culturali indispensabili alla sua dignità ed al libero sviluppo della sua personalità*"; ed agli articoli 8 e 9 della Carta europea dei diritti dell'uomo (CEDU), che sanciscono il diritto al rispetto della vita privata e familiare ed alle libertà di pensiero, di coscienza e di religione, garantendo a chiunque la tutela della propria sfera privata.

[24] O. POLLICINO, *Costituzionalismo, privacy e neurodiritti,* in MediaLaw, 2.04.2021; E. BELISARIO, *Neurodiritti:nuovi diritti o diritti già esistenti?,* in Federalismi.it, 15.03.202; C. SLIGHTHART, S. BUBLITZ, S. ALEGRE, *Neurotechnology: we new laws, not new rights*, in Nature, 9.08.2023.

opposizione per sottrarsi alle attività di personalizzazione dei messaggi inoltrati dall'AI Companion e, di conseguenza, alla persuasione ed alla manipolazione eventualmente in corso.

Per offrire spunti concreti si ricorda che Meta Platforms Technologies Ireland Limited, titolare dei social network Facebook e Instagram, ha reso noto, già dall'agosto del 2023, la strategia di business di utilizzare i dati personali presenti sulle piattaforme relativi ad account pubblici e quelli relativi alle interazioni dell'utente con il modello di intelligenza artificiale Meta AI a fini di addestramento dello stesso. Ciò significa che, in assenza dell'esercizio del diritto di opposizione (opt-out) da parte dell'interessato (utente o non utente) i dati condivisi pubblicamente sulla piattaforma (public audience) da utenti di età pari o superiore a 18 anni, come post e commenti, ed i dati generati dall'utilizzo dei modelli intelligenti di Meta potranno essere utilizzati per allenare la loro capacità di "comprensione" e per perfezionarli[25].

Dalla privacy policy della Società si apprende che il trattamento di tali dati avviene sulla base giuridica del legittimo interesse di cui all'articolo 6, par.1, lettera f) del Regolamento e che la stessa, in sede di Legitimate interest assessment (nota come L.I.A.), ha eseguito una valutazione preventiva che tiene conto dei diritti e delle libertà fondamentali degli interessati.

Dunque, sono riconosciuti all'utente delle piattaforme gestite da Meta da un lato il diritto di essere informato, preventivamente, delle attività di trattamento, *in primis* attraverso le informazioni di cui al Capo III del Regolamento e, dall'altro, il diritto di opporsi a tali attività.

Inoltre, nel caso di specie, vale la pena sottolineare che il diritto di opposizione, coerentemente a quanto stabilito nella Opinion recentemente adottata dall'EDPB ai sensi dell'articolo 64, par.2, del Regolamento [26], è incondizionato, per cui, diversamente da quanto previsto dalla normativa stessa, è sempre accolto dal titolare del trattamento.

Risulta evidente che la normativa in argomento trova applicazione in quanto i dati coinvolti nelle attività di training dei modelli di IA, a cui l'utente si relaziona sottoforma di chatbot, sono a tutti gli effetti riconducibili alla nozione di "dato personale" contemplata dall'articolo 4 del Regolamento.

Ciononostante, esaminando la fattispecie nella sua interezza, si nota che, spesso, le conversazioni tra agente virtuale e utente si snodano talmente assiduamente e in un clima così confidenziale ed intimo che i dati coinvolti potrebbero essere idonei a rivelare informazioni uniche non comuni, capaci di descrivere i processi cerebrali associati alla personalità, all'umore, ai pensieri o ai sentimenti.

Al momento i principali modelli sembrano in grado di identificare il profilo psicologico dell'utente e di analizzare il linguaggio per simulare una mappa mentale utile ai fini della personalizzazione dell'interazione [27], ma non di tracciare un profilo neurale in senso stretto. Se però questo dovesse verificarsi,

---

[25] GPDP, *Intelligenza artificiale, Garante privacy: Da fine maggio Meta addestrerà i suoi sistemi utilizzando i dati personali degli utenti che non si saranno opposti,* Comunicato stampa, 29.04.2025. Consultabile sul sito: www.garanteprivacy.it.

[26] EDPB, *Parere 28/2024 su taluni aspetti relativi alla protezione dei dati ai fini del trattamento dei dati personali nel contesto dei modelli di IA*, 17.12.2024, Consultabile sul sito: https://www.edpb.europa.eu/.

[27] M. PANETTO, *Amicizie virtuali, emozioni reali: AI companion tra rischi e benefici*, in Live Università di Padova, 11.08.2024, Consultabile al sito: https://ilbolive.unipd.it.

se dunque in un prossimo futuro dovesse essere possibile per gli AI Companion acquisire informazioni neurali capaci di rivelare ciò che un uomo pensa, il grado di sensibilità ed il valore di tali informazioni le collocherebbe in una posizione del tutto nuova rispetto alle categorie regolate e codificate dal legislatore europeo e lo standard di protezione attuale non risulterebbe più in grado di rispondere alle emergenti esigenze di tutela[28].

In tale contesto rilevano altresì le nozione di profilazione descritta dall'articolo 4, paragrafo 1), numero 4, del Regolamento come "qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica". Tali attività, infatti, se non correttamente eseguite e fortemente limitate, potrebbero esporre la persona al rischio che i suoi dati personali, segnatamente le conversazioni intrattenute con l'agente in un ambiente apparentemente neutro e sicuro, possano essere utilizzati dallo stesso fornitore del servizio di chatbot per condizionare l'utente, o anche per cederle a soggetti terzi per altri fini, tra cui quelli di marketing.

De iure condendo pare possibile affermare che la collocazione dei dati neurali tra i dati particolari ai sensi dell'articolo 9 del Regolamento, indipendentemente dalla circostanza per cui i processi emotivi e cognitivi a cui si riferiscono siano riconducibili a specifiche sfere dell'identità personale, come la vita sessuale, il credo religioso e l'appartenenza politica, consentirebbe l'applicazione di un regime di maggior tutela, segnatamente di un divieto generale di trattamento superabile solo in presenza di talune eccezioni.

L'ampliamento della portata dell'articolo 9 del Regolamento potrebbe comportare un primo passo verso una tutela che si rende più che mai necessaria in uno scenario del genere, considerato che, come è stato efficacemente osservato[29], "quale significato avrebbe la tutela dell'intangibilità della sfera privata, in ogni sua articolazione, se, poi, i pensieri fossero leggibili e venisse così negata la riservatezza di quei thoughts, emotions, sentiments and sensations che già Warren e Brandeis, nel 1890, indicavano come fondamento essenziale del right to privacy? Può dirsi realmente libertà se l'uomo, mediante la tecnica, diviene osservatore delle più intime percezioni, aspirazioni, volontà altrui (e persino proprie, se ignote)? "[30].

6.2 La disciplina dei contratti e quella consumeristica.

Il condizionamento del processo di formazione della volontà acquista rilievo anche nel contesto in cui la persona assume la veste di consumatore, vale a dire qualora un utente venga condizionato dalla relazione amicale instaurata con un AI Companion per fini consumeristici.

---

[28] D. HALLINAN, P. SCHUTZ, M. FRIEDEWALD E P.DE HERT, *Neurodata and Neuroprivacy: Data Protection Outdated?,* in Surveillance & Society, 1.01.2014.
[29] P. STANZIONE, *Giornata europea della protezione dei dati 2021*, in www.gpgp.it, Roma, 28.01.2021.
[30] O. POLLICINO, *Costituzionalismo, privacy e Neurodiritti*, in MediaLaw, 2.04.2021.

La comunità scientifica[31] ha evidenziato come, in applicazione del criterio ermeneutico della compatibilità, al fine di ottenere tutela rispetto alle diverse forme di marketing, sarebbe possibile ricorrere agli istituti delle invalidità civilistiche della nullità e dell'annullabilità, ma anche come tali rimedi siano ontologicamente inadatti. Sarebbe infatti opportuno intervenire con provvedimenti specifici che intervengano a tutela della libertà di autodeterminazione del consumatore contro le tecniche di neuromarketing[32] acutamente affinate tramite i dati raccolti durante l'interazione con un AI Companion, come l'eye tracking, l'analisi delle espressioni facciali, i pattern linguistici, i feedback impliciti ed il tono emotivo dell'utente, da cui potrebbe essere dedotto il grado di engagement dell'utente, le sue preferenze ed il suo grado di entusiasmo.

Il codice del 1942, pensato ed adottato in una epoca in cui la promozione delle attività commerciali non contemplava né la commercializzazione di massa né la personalizzazione delle comunicazioni commerciali, appare uno strumento non perfettamente idoneo alla tutela del consumatore della dimensione digitale.

Ai sensi dell'articolo 1418 del Codice civile il contratto è nullo quando è contrario a norme imperative, salvo che la legge disponga diversamente.

Pertanto, ammesso che il divieto di pratiche commerciali scorrette stabilito dal Codice del consumo possa essere interpretato come una norma imperativa, quindi inderogabile, sarebbe possibile per il consumatore, vittima di un indebito condizionamento da parte di un AI Companion, far accertare *sine die*[33] che l'accordo negoziale sia improduttivo di effetti.

Il comma 2 della stessa norma prevede che il contratto sia nullo laddove manchi un elemento essenziale, sicché se si considerasse assente la volontà del contraente in ragione della manomissione della sua capacità di scelta da parte dell'agente intelligente il contratto potrebbe essere ritenuto improduttivo di effetti.

Nel corso degli anni la dottrina ha affiancato alle tipiche figure di invalidità ipotesi di nullità peculiari, caratterizzate dall'intento di offrire maggior tutela alla parte del negozio presumibilmente più debole. La nullità in questione è definita a legittimazione relativa in quanto è rilevabile solo dalla parte che goda di minor potere contrattuale, oltre che d'ufficio dal giudice. Lo squilibrio che si intende prevenire o elidere tramite la disciplina delle nullità relative è, normalmente, di natura contrattuale e solo eccezionalmente economica[34]. Di regola, dunque, ai fini della valutazione di una manipolazione della volontà del consumatore ad opera di tecniche di neuromarketing, non dovrebbe assumere rilievo la congruità del corrispettivo pagato rispetto al valore del bene, ma l'equilibrio normativo del contratto, vale a dire gli obblighi ed i diritti che discendono dallo stesso.

---

[31] A. PANICHELLA, *La tutela del consumatore alla sfida del neuromarketing*, in Federalismi.it, 2024.

[32] M. LINDSTROM, *Neuroeconomia, neuromarketing. Attività cerebrale e comportamenti d'acquisto*, Apogeo, 2009.

[33] Articolo 1422 del Codice Civile.

[34] L'articolo 33 del Codice del Consumo, nel definire la clausola vessatoria, afferma espressamente che essa è tale quando provoca un significativo squilibrio tra i diritti e gli obblighi derivanti dal contratto.

Dal punto di vista degli effetti, la nullità di protezione, virtuale e relativa, ha un effetto manutentivo, per cui nel caso in cui l'invalidità venisse accertata non sarebbe sanzionato l'intero assetto negoziale, ma le singole parti del contratto ritenute pregiudizievoli per il contraente che ha subito la scorrettezza del professionista.

Il rimedio della nullità relativa può operare solo nella misura in cui il condizionamento del consumatore non abbia inciso sulla sua volontà a contrarre, ma su aspetti specifici del negozio che, in assenza dell'attività di neuromarketing, sarebbero stati definiti in modo diverso dal contraente.

Le invasive pratiche di neuromarketing potrebbero essere fronteggiate anche attraverso il ricorso all'istituto dell'annullabilità ed il contratto risultare invalido *ex* articolo 1439 del Codice civile, dimostrando che l'elemento perturbatore ha condizionato la volontà del contraente debole dolosamente ed in modo determinante; ciò anche laddove le attività dolose provengano da un soggetto terzo e la parte che beneficia degli artifici e dei raggiri ne sia a conoscenza. Potrebbe essere il caso in cui un service provider di servizi di AI Companion condizioni la volontà dell'utente per conto di un terzo. L'efficacia di tale rimedio civilistico però potrebbe essere minata dal fatto che il raggiro è individuato solo atipicamente dal legislatore, pertanto l'interprete per dargli forma e capire se una pratica di neuromarketing abbia effettivamente portato l'utente consumatore ad assumere scelte diverse da quelle che avrebbe voluto, dovrebbe ricorrere alla disciplina dei vizi della volontà, che risultano particolarmente difficili da dimostrate a fronte di accurate strategie di marketing personalizzate.

Ne risulta che i rimedi offerti dalla disciplina civilistica, pensati per ingerenze occasionali nel processo di formazione della decisione dell'acquirente, pur offrendo significativi spunti, una volta applicati a casi concreti, si rivelano fragili rispetto alle solide strategie di vendita elaborate attraverso dati raccolti durante l'interazione con un AI Companion. Pertanto, come evidenziato da attenta dottrina, "di fronte alle nuove insidie e pericoli per la persona non è percorribile la strada della verifica, di volta in volta e sul piano del singolo contratto, della possibile sussistenza di un occasionale vizio del consenso nel processo formativo della volontà." [35] . Sarebbe dunque auspicabile un intervento legislativo che tenga conto del mutamento radicale che è intervenuto e che sta intervenendo nel mercato del consumo.

Un'altra strada percorribile per tutelare le scelte dell'utente in ambito consumeristico rispetto alle efficaci attività di marketing elaborate sulla base dello studio dell'utente è rappresentata dal ricorso alle norme dettate dal Codice del consumo in materia di pratiche commerciali sleali, che trovano applicazione, trasversalmente, a tutti i contratti che prevedano uno scambio tra un consumatore ed un professionista ed hanno l'obiettivo ambivalente di tutelare la parte debole del rapporto obbligatorio ed assicurare il corretto funzionamento del mercato.

L'articolo 21, comma 1, del Codice del consumo considera ingannevoli le azioni che in qualsiasi modo inducano, o sono idonee ad indurre, il consumatore medio ad assumere una decisione di natura commerciale, che altrimenti non

[35] L. TAFARO, *Neuromarketing e tutela del consenso,* Napoli, 2018.

avrebbe preso, riguardo ad uno o più dei seguenti elementi: l'esistenza o la natura del prodotto; le caratteristiche principali del prodotto; la portata degli impegni del professionista, i motivi della pratica commerciale e la natura del processo di vendita; il prezzo o il modo in cui questo è calcolato o l'esistenza di uno specifico vantaggio; la necessità di un intervento di manutenzione, ricambio, sostituzione o riparazione; la natura, le qualifiche e i diritti del professionista o del suo agente; i diritti del consumatore.

Al ricorrere di determinati presupposti le attività di marketing che si basano sullo studio e la disamina delle attività dell'utente in un ambiente fintamente protetto e confidenziale potrebbero essere considerate anche pratiche commerciali aggressive, *ex* articolo 24 del Codice del consumo[36], in quanto capaci di generare una coercizione o un indebito condizionamento del consumatore tali da limitare considerevolmente la sua libertà di scelta o di comportamento in relazione ad un prodotto e, pertanto, ad indurlo ad assumere decisioni che non avrebbe altrimenti preso.

Tuttavia, l'inadeguatezza già riscontrata commentando le norme civilistiche emerge anche con riferimento alla cornice normativa in argomento[37], che impiega un parametro, quale quello del consumatore medio[38], evidentemente inadatto ai fini della valutazione delle pratiche commerciali sleali. Infatti, nell'infosfera, ove la vulnerabilità è generalizzata[39] e chiunque può essere vittima di trigger che agiscono nel subconscio, la valutazione dell'efficacia delle strategie di marketing dovrebbe avvenire in considerazione delle peculiarità della fattispecie concreta, esulando da parametri astratti e considerando le reali capacità delle strategie commerciali di condizionare o, addirittura, manipolare indebitamente la capacità del consumatore di orientarsi in modo libero nell'ecosistema commerciale.

Come nel caso della normativa in materia di dati personali, anche in questo ambito sarebbe quindi opportuno un intervento legislativo *ad hoc,* che introduca parametri flessibili ed abbandoni il criterio del consumatore medio.

6.3 La strategia digitale europea.

L'interazione con uno dei modelli di IA più avanzati disponibili sul mercato europeo si apre con un messaggio che guida l'utente tra potenzialità e limiti dell'intelligenza artificiale: "ogni conversazione con un personaggio è accompagnata da un avviso che informa gli utenti che si tratta di un chatbot basato sull'intelligenza artificiale e non di una persona reale. Ciò che viene espresso è da considerarsi una finzione. Non considerare ciò che viene espresso come una verità o un consiglio".

---

[36] A. PANICHELLA, *La tutela del consumatore alla sfida del neuromarketing*, in Federalismi.it, 15.03.2024.
[37] M. ADDIS, *Le pratiche commerciali sleali e le risorse di fiducia delle imprese: aspetti positivi e questioni irrisolte,* Milano, 2007.
[38] L'articolo 14 del Codice del consumo recita: "È considerata aggressiva una pratica commerciale che, nella fattispecie concreta, tenuto conto di tutte le caratteristiche e circostanze del caso, mediante molestie, coercizione, compreso il ricorso alla forza fisica o indebito condizionamento, limita o è idonea a limitare considerevolmente la libertà di scelta o di comportamento del consumatore medio in relazione al prodotto e, pertanto, lo induce o è idonea ad indurlo ad assumere una decisione di natura commerciale che non avrebbe altrimenti preso".
[39] Corte Giust., 16.07.1998, C.210/96, *Gut Springeheide GmbH, RudolfTusky c. Oberkreisdirektor des Kreises Steinfurt*, Consultabile al sito: www.curia.europa.eu.

Il disclaimer risponde all'obbligo di trasparenza previsto dall'AI Act per i fornitori di determinati sistemi di IA. In particolare, l'articolo 50 del Regolamento impone ai fornitori di integrare nei sistemi soluzioni tecniche che consentano agli output di essere rilevabili come generati o manipolati da un modello di IA, al fine di preservare l'integrità e la fiducia nell'ecosistema dell'informazione, e di fronteggiare i rischi di una cattiva informazione, di una manipolazione, e di un inganno su vasta scala.

Parimenti il Regolamento UE 2022/2065 relativo al mercato unico dei servizi digitali (DSA) mira ad accrescere la consapevolezza degli utenti, stabilendo obblighi di trasparenza a carico delle piattaforme di grandi dimensioni[40]. L'articolo 26 del DSA regola le attività pubblicitarie sulle piattaforme online, ampliando le prerogative dei destinatari dei servizi digitali attribuendo loro il diritto di essere messi in grado di identificare in modo chiaro, conciso, inequivocabile ed in tempo reale che "l'informazione costituisce una pubblicità; la persona fisica o giuridica per conto della quale viene presentata la pubblicità; la persona fisica o giuridica che paga per la pubblicità; le informazioni relative ai parametri utilizzati per determinare il destinatario al quale viene presentata la pubblicità e, laddove applicabile, alle modalità di modifica di detti parametri".

Inoltre, l'articolo 27 del DSA mira a riequilibrare l'asimmetria informativa che caratterizza i rapporti business to consumer (B2C), attraverso il riconoscimento agli utenti del diritto di conoscere i criteri su cui si basa il funzionamento dei sistemi di raccomandazione dei contenuti e le ragioni per cui vengono suggerite determinate informazioni. Qualora siano disponibili diversi parametri per modulare il funzionamento dei sistemi di raccomandazione, i fornitori di piattaforme online devono mettere a disposizione dell'utente una funzionalità che consenta loro di selezionare e modificare, in qualsiasi momento, i parametri impostati di default, in modo da determinare autonomamente l'ordine dei contenuti che vengono mostrati.

Tali previsioni rispondono all'intento del legislatore di promuovere buone prassi e di rendere meno opaco il funzionamento dell'ecosistema pubblicitario, consentendo all'interessato di conoscere ed affrontare le pratiche discriminatorie che si insidiano nella realtà digitale.

La disciplina introdotta dall'AI Act si distingue dalle norme sopramenzionate in quanto caratterizzata da una portata significativamente più ampia e non rivolta agli accordi negoziali.

In particolare, l'articolo 5 dell'AI Act mira a prevenire qualsiasi forma di dominio psicologico esercitato sulla persona mediante modelli di intelligenza artificiale. A tal fine esso vieta l'impiego di sistemi che sfruttino le vulnerabilità individuali, nonché l'utilizzo di tecniche subliminali, manipolative o ingannevoli, qualificando il rischio derivante da tali pratiche come inaccettabile e, conseguentemente, proibendone l'immissione sul mercato, la messa in servizio e l'utilizzo[41].

La pratica manipolativa, per assumere rilievo, deve corrispondere alla puntuale descrizione di cui all'articolo 5, paragrafo 1, lett. a), dell'AI ACT

[40] La stessa previsione normativa prevede altresì che l'utente sia messo nelle condizioni di comprendere la natura pubblicitaria delle informazioni visualizzate e di identificare la persona fisica o giuridica per conto della quale viene visualizzata la pubblicità.

[41] Articolo 5, par. 1, lett. a) del Regolamento UE 2024/1689 in materia di intelligenza artificiale.

segnatamente deve: comportare l'adozione da parte dell'interessato di una decisione che altrimenti non avrebbe assunto, essere orientata ad uno scopo illecito, e arrecare un danno, fisico o psicologico, alla persona che ne sia vittima o ad un altro individuo. Spetta all'interprete valutare la sussistenza dei presupposti che rendano operativo il divieto e distinguere tali da quelle meramente persuasive (c.d. nudging), che, pur influenzando la volontà dell'individuo, sono ritenute ammissibili in quanto non idonee a compromettere in maniera totalizzante le capacità di rappresentazione e di autodeterminazione della persona.

La scelta del legislatore comunitario si inserisce all'interno di un quadro regolatorio delle intelligenze artificiali improntato ad un approccio antropocentrico e precauzionale, nonché al principio etico dell'autodeterminazione dell'individuo.

Eppure, la marcata variabilità che contraddistingue la dimensione psicologica individuale, unitamente all'assenza di un modello condiviso di riferimento e alla difficoltà di distinguere le pratiche manipolative dalle semplici forme di persuasione, rende particolarmente arduo dimostrare di essere stati vittima di condotte manipolative illecite. Tale prova risulta più agevole soltanto nei casi in cui sia possibile individuare un'indicazione esplicita, impartita dal programmatore o dallo sviluppatore del modello, volta a sfruttare deliberatamente le vulnerabilità dell'interessato.

Dal tenore letterale della norma sembrerebbe emergere che non sia sufficiente dimostrare il mero danno-evento, inteso come alterazione concreta del comportamento causata da un'attività manipolatoria, ma che sia invece richiesta la prova di un danno "significativo", pur non essendo stati esplicitati i criteri in base ai quali tale significatività debba essere valutata.

D'altro canto, proprio la natura artificiale delle interazioni manipolative dovrebbe poter consentire la misurabilità di tali fenomeni.

Secondo studi di settore il ricorso a tecniche psicometriche potrebbe consentire di quantificare e comprendere l'impatto dell'IA sulla sfera neurocognitiva della persona. Gli esperti della materia individuano nello State -Trait Sense of Self Inventory (st-SoSI) uno strumento psicometrico prezioso per misurare in modo multidimensionale il senso del sé: il primo inventario che consente di distinguere tra aspetti stabili e dinamici della personalità, offrendo una lettura fine delle interazioni tra identità, agentività e tecnologia. Il st-SoSI risulterebbe particolarmente utile per comprendere come il senso di sé si modifichi in relazione al contesto, alle vulnerabilità individuali, ed alle interazioni con dispositivi tecnologici intrusivi, e per quantificare la capacità di questi ultimi di influenzare l'individuo[42].

[42] S.D. PLINIO, *Neurodiritti tra virtuale e reale. Navigare il sé digitale: implicazioni delle neurotecnolgie sull'autonomia cognitiva e la consapevolezza dei diritti*, in Federalismi, 16.03.2024; S.D. PLINIO, *The state-trait sense of self inventory: A psychometric study of self-experience and its relation to psychosis-like manifestations,* IRIS, 1.01.2024, Consultabile su: https://ricerca.unich.it/

7. L'educazione digitale.

Dal blog di OpenAI, una dei fornitori più importanti dei servizi di IA, si apprende che la Società sia impegnata a "migliorare i modelli e a sviluppare strumenti che aiutino a individuare sintomi di disagio mentale o emotivo (seppur rari), così che ChatGPT possa rispondere in modo adeguato e, se necessario, indirizzare gli utenti a consultare risorse basate su evidenze scientifiche". In tale prospettiva da qualche mese la Società prevede che durante le sessioni più lunghe, l'utente riceverà dei promemoria che lo invitano a fare una pausa. Inoltre, la Società informa che sta "collaborando con esperti per migliorare il modo in cui ChatGPT risponde nei momenti critici, ad esempio quando una persona mostra segni di disagio mentale o emotivo" attraverso la collaborazione "con oltre 90 medici (tra cui psichiatri, pediatri e medici di base) provenienti da più di 30 Paesi, per creare criteri personalizzati in grado di valutare conversazioni complesse e articolate"[43].

Aumentare il grado di vigilanza contro le intrusioni nella sfera dei pensieri e delle pulsioni e promuovere l'esercizio dello spirito critico e di giudizio appaiono i principali strumenti a disposizione della persona per arginare i rischi derivanti dall'avanzamento del digitale nelle sue diverse declinazioni.

Ad esempio una maggiore cura e riservatezza dei propri dati personali potrebbe costituire parte della soluzione, atteso che i dati forniti direttamente ed indirettamente dagli interessati, quelli pubblicamente disponibili e quelli inferenziali sono fondamentali ai fini dell'esercizio di attività di condizionamento, e che, pertanto, la minore disponibilità delle informazioni di carattere personale renderebbe più difficile la costruzione di meccanismi di persuasione efficaci.

Dunque, oltre all'adozione di interventi legislativi mirati, la soluzione sembra risiedere nella sensibilizzazione delle persone[44] e nella promozione e nello sviluppo del pensiero critico, sia sul piano individuale che collettivo, attraverso iniziative educative. L'educazione, in tal senso, si configurerebbe come uno strumento di resistenza all'intrusione del digitale, capace di rafforzare l'autonomia del pensiero e la libertà di scelta.

L'auspicio è che tale approccio possa esercitare un'influenza positiva anche sul dibattito politico e sul processo democratico[45].

La valorizzazione della letteratura digitale permetterebbe alle persone di rapportarsi a sistemi di IA nella consapevolezza che si tratta di mere tecnologie che restituiscono output statistici, di riconoscerle quindi per quel che realmente sono, di comprenderne la logica e di conoscere i diritti attribuiti dall'ordinamento per reagire e difendersi.

Siamo il risultato delle nostre decisioni e viviamo entro lo spazio che esse delineano. Quando questo spazio si restringe perché un algoritmo decide al posto nostro, o un sistema di IA ci orienta verso una determinata scelta, rinunciamo all'esercizio autentico della libertà di giudizio, affidandoci a una

[43] OpenaAI. Blog Consultabile sul sito https://openai.com/it-IT/index/how-we're-optimizing-chatgpt/

[44] N.A. SMUHA, M. DE KETELAERE, M. COECKELBERGH, P. DEWITTE, Y. POULLET, *Open Letter: We are not ready for manipulative AI – urgent need for action,* 31.03.2023, in Ku Leuven – AI Summer School, Consultabile sul sito: https://www.law.kuleuven.be/.

[45] L. FLORIDI, *Hypersuassion – ON AI'S persuasive power and how to deal with*, Bologna, 3.05.2024.

versione arbitraria e semplificata dello spettro delle possibilità, modellata dal calcolo digitale[46].

L'intelligenza artificiale è stata spesso paragonata al fuoco: una forza primordiale che, se sapientemente governata, può offrire calore, nutrimento e protezione, ma che, se lasciata agire senza controllo, può trasformarsi in un agente distruttivo, capace di incenerire tutto ciò che incontra sul proprio cammino.

Siamo dinanzi ad una potenza immensa e, proprio come con il fuoco, è indispensabile comprenderne i limiti e imparare a controllarla.

8. Conclusioni.

La panoramica delle norme che direttamente ed indirettamente governano la materia suggerisce che, ad oggi, in Italia e in Europa, gli aspetti legati all'influenza e manipolazione delle persone per fini economici e commerciali siano i più approfonditi e regolati, seppur con numerosi vulnus che suggeriscono degli interventi.

Al contrario risulta ancora limitata la protezione giuridica rispetto ai condizionamenti che incidono sulla sfera individuale in contesti diversi da quello strettamente consumeristico. In tale prospettiva emerge con chiarezza l'esigenza di affiancare all'intervento normativo strategie alternative capaci di incidere profondamente sulla coscienza individuale.

Percorsi di educazione critica, di valorizzazione delle relazioni interpersonali e la riscoperta del senso di comunità e dei legami umani, che costituiscono il fondamento della convivenza civile, sono efficaci strumenti per il contrasto all'influenza pervasiva dei meccanismi digitali e la salvaguardia dell'autonomia della persona[47].

Dinanzi alle notizie tragiche riportate dalla stampa[48] come quelle di una minore uccisa dal proprio fidanzato, che la perseguitava, e che giorni prima dell'omicidio si era rivolta a ChatGpt per chiedere se la storia che stava vivendo fosse tossica, capiamo come, probabilmente, ciò su cui si deve intervenire non sono le norme giuridiche ma le relazioni umane, il dialogo tra amici, figli, insegnanti e genitori. Dunque, bisognerebbe riacquisire la capacità di mostrarsi indifesi e fragili alle persone care, non frenando le proprie confidenze per paura di un giudizio, ma riuscendo a mostrarsi imperfetti, fragili e indifesi.

Il mantra dell'infallibilità su cui è costruita la società del nostro tempo dovrebbe essere sostituito da quello del dialogo e della condivisione, restituendo all'uomo ciò che mai la macchina potrà sottrargli: la capacità di instaurare relazioni reali e di tendere una mano a chi ne ha bisogno.

---

[46] M. CHIRIATTI, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Luiss University Press, Roma.
[47] S.D. PLINIO, *Neurodiritti tra virtuale e reale. Navigare il sé digitale: implicazioni delle neurotecnologie sull'autonomia cognitiva e la consapevolezza dei diritti*, in Federalismi, 2024.
[48] U. MILANO, *Aurora Tila, le domande della 13enne a ChatGPT sull'ex fidanzato: «È un amore tossico?* La Stampa, 20.06.2025, Consultabile sul sito: www.lastampa.it.