lacittadinanzaeuropea online
Rivista di studi e documentazione sull'integrazione europea
ISSN 2785-4183
lceonline 2/2025 (*www.lceonline.eu*)

L'impatto dell'Intelligenza Artificiale sui minori (*)

Luis Fernando MARTÍNEZ QUEVEDO*

Sommario: 1. Introduzione **2.** Minori: utenti e destinatari maggiormente vulnerabili **3.** Minori e Intelligenza Artificiale: rischi specifici *3.1.* Il mondo dei giocattoli intelligenti *3.2.* Algoritmi che raccomandano contenuti sui social network *3.3.* Deepfakes **4.** Conclusioni

1. Introduzione

Dalla fine del secolo scorso, siamo immersi in un processo di costante progresso nel campo digitale. Le nuove tecnologie sono diventate una componente imprescindibile della nostra società, trascendendo inesorabilmente gli ambiti per cui erano state originariamente concepite. Secondo Agustinoy Guilayn e Monclús Ruiz, le reti sociali, ad esempio, sono nate come un metodo per stabilire contatti basato sulla teoria dei "sei gradi"[1]; attualmente, tuttavia, rappresentano qualcosa di molto più profondo, in quanto costituiscono un contesto di comunicazione continua, di condivisione e di ricezione di informazioni di varia natura e rilevanza, con i profili degli utenti che diventano un'ulteriore sfaccettatura della loro identità.

Questo progresso ci colloca in un contesto che cambia di continuo e incessantemente. Non comprendiamo più la vita come veniva concepita all'inizio del XXI secolo. Una delle caratteristiche principali dell'epoca in cui viviamo è la velocità: i progressi sono costanti e la loro diffusione è evidente. L'offerta di sempre nuovi dispositivi con sempre nuove funzionalità avviene a un ritmo tale da non avere quasi il tempo di assimilarli appieno. In questo senso, è frequente che, proprio quando iniziamo a familiarizzare con una nuova tecnologia, ne emerga un'altra ancora più avanzata, che in parte condanna la precedente all'obsolescenza.

Possiamo parlare dell'intelligenza artificiale in termini analoghi. Nel 2022 è stato adottato il Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio, del 19 ottobre 2022, relativo a un mercato unico dei servizi digitali e recante modifica della direttiva 2000/31/CE (di seguito, DSA [Digital Services Act]). Si tratta di un regolamento attraverso il quale l'UE intende rafforzare la tutela degli utenti dei servizi digitali, affrontando complesse questioni legate a tecnologie che, pur appartenendo a una fase precedente dell'evoluzione digitale, continuano (e continueranno) a far parte della nostra vita quotidiana.

A esso si affianca il Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13

(*) Traduzione dal castigliano di Luca Maldarella.

* Contratado Predoctoral-FPU, Universidad de Granada, Departamento de Derecho Constitucional.

[1] A. Agustinoy Guilayn e J. Monclús Ruiz, *Aspectos legales de las redes sociales,* Barcelona, 2016, p. 19: "el origen del concepto de red social puede encontrarse en la teoría denominada de 'los seis grados' enunciada originalmente por Karinthy en 1929. De acuerdo con la misma, cualquier persona puede interactuar con cualquier otra persona en el mundo a través de un máximo de seis conexiones". Come osserva Davara Fernández de Marcos, "lo cierto es que la primera red social, tal y como ahora la concebimos, nace en 1995 bajo la denominación 'classmates', con el objetivo de poner en contacto a personas de un mismo colegio, una vez dejaran el mismo", aggiungendo che non ha avuto il successo sperato; si veda a questo proposito L. Davara Fernández de Marcos, *Régimen jurídico de las redes sociales*, in V. Cazurro Barahona (Dir.), *Derecho Digital. Perspectiva interdisciplinar,* Barcelona, 2017.

giugno 2024, che stabilisce norme armonizzate in materia di intelligenza artificiale e modifica numerosi regolamenti e direttive europee (di seguito, Regolamento sull'intelligenza artificiale – RIA), nonché la Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto (di seguito, Convenzione quadro-IA), che mirano a rispondere alle sfide poste dalle tecnologie più recenti, i cui limiti restano tuttora indefiniti.

In questo contesto, sia le istituzioni sia – naturalmente – gli operatori e gli studiosi stanno cercando di fare luce su una realtà ancora poco definita, che avvolge un futuro che ritenevamo più lontano, ma che è ormai divenuto presente. Numerose sono le sfide che il mondo del diritto è chiamato ad affrontare: l'assunzione di responsabilità, i limiti etici di alcune funzionalità e le conseguenze che l'implementazione di determinate tecnologie può comportare in settori quali l'occupazione, la salute o l'istruzione che ne sono esempi significativi.

È in questo contesto che si colloca il presente contributo. Il nostro obiettivo è rivolgere l'attenzione alla condizione dei minori, gruppo che merita sempre una considerazione specifica, e alle modalità per cui l'intelligenza artificiale possa rappresentare un rischio particolarmente rilevante a loro danno. In sintesi, intendiamo analizzare le minacce che, in questo ambito specifico di studio, accompagnano l'uso dell'intelligenza artificiale da parte dei soggetti in età evolutiva.

2. I minori: utenti e destinatari maggiormente vulnerabili

Non c'è dubbio che la rivoluzione digitale abbia introdotto profondi cambiamenti nella nostra vita quotidiana. Oggi assistiamo all'emergere di una nuova tecnologia dirompente, l'intelligenza artificiale, che rappresenta al tempo stesso un'evoluzione della fase precedente e una novità radicale, pur essendo il suo concetto apparso già a metà del XX secolo[2].

Questa irruzione giunge in un momento in cui siamo ancora impegnati a interrogarci su come prevenire i danni a cui sono esposti gli utenti del web 2.0 e dei social network, su come attribuire le responsabilità e su come risarcire chi subisce un pregiudizio. L'arrivo dell'intelligenza artificiale può dunque apparire, per certi versi, precoce. Esistono diversi argomenti a sostegno di questa affermazione, tra cui due spiccano in modo particolare, l'uno strettamente conseguente all'altro. In primo luogo, non conosciamo ancora i limiti dell'intelligenza artificiale, il suo margine di miglioramento nel breve periodo, né la rapidità con cui queste tecnologie si svilupperanno. In secondo luogo, proprio per questa incertezza, non siamo ancora in grado di comprendere realisticamente a cosa siano esposte le persone a causa del suo utilizzo, quali diritti siano potenzialmente a rischio e come gestire situazioni di questo tipo.

Tuttavia, il fatto di non conoscere ancora i limiti dell'intelligenza artificiale non implica che non si possa analizzare ciò che è già presente, ossia le funzionalità e le applicazioni basate su questa tecnologia che risultano già ben consolidate.

Dal punto di vista del diritto e in particolare del diritto costituzionale, una delle questioni principali, riguarda la minaccia cui sono sottoposti i diritti fondamentali degli utenti o dei destinatari di tali strumenti[3]. La dottrina ha concentrato in gran parte la propria attenzione sull'impatto del fenomeno sul diritto alla protezione dei dati personali, in linea con l'approccio adottato da istituzioni come

[2] Il termine "intelligenza artificiale" è apparso a metà del secolo scorso, grazie a diversi studi condotti da ricercatori come Alan Touring, Norbert Wiener e John von Neumann. Tuttavia, l'origine canonica dell'espressione risale al convegno scientifico tenutosi al Darmouth College nel New Hampshire, intitolato "Summer Research Project in Artificial intelligence" e organizzato da John McCarthy, Claude Shannon e Marvin Minsky. Cfr. P. Meseguer González e R. López de Mántaras Badía, *Inteligencia artificial,* Madrid, p. 17 ss.

[3] Cfr., tra gli altri, L. Gómez Abeja, *Inteligencia artificial y derechos fundamentales*, in F. H. Llano Alonso (Dir.), *Inteligencia artificial y Filosofía del Derecho,* Murcia, 2022, p. 93: "existen otros derechos incididos a consecuencia del uso del *big data* y del desarrollo de la inteligencia artificial y sus avanzados algoritmos para la toma de decisiones, como el derecho a la tutela judicial [...] el derecho a la libertad de información [...] el derecho al sufragio [...] o el derecho de acceso a la información pública".

l'Unione europea, fortemente impegnata nella regolamentazione delle nuove tecnologie proprio a partire da questo diritto. Un chiaro esempio è rappresentato dal Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali e alla libera circolazione di tali dati, che ha abrogato la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati, di seguito GDPR).

Un ulteriore aspetto da tenere in considerazione è che l'intelligenza artificiale può generare rischi per la stessa democrazia, come dimostra il noto caso di Cambridge Analytica, società che ha raccolto un'ingente quantità di dati dagli utenti di Facebook, successivamente utilizzati a fini di propaganda politica a vantaggio di Donald Trump nelle elezioni presidenziali statunitensi del 2016.

È in questo contesto che si giustifica pienamente il titolo del presente contributo. Vale a dire: considerato che le applicazioni fondate sull'intelligenza artificiale – basate su big data, machine learning, deepfake, ecc. – sono già capaci di scuotere le fondamenta stesse delle nostre democrazie, e noi adulti ignoriamo ancora il pieno potenziale di questi strumenti e, in molte occasioni, risultiamo esposti a violazioni dei nostri diritti fondamentali, pur disponendo di maggiori risorse e strumenti per reagire; allora, quale può essere il margine di esposizione quando gli utenti o i destinatari sono i minori?

Si potrebbe affermare, in proposito, che una persona adulta con una conoscenza media dei benefici e dei rischi associati alle nuove tecnologie è, in una certa misura, vulnerabile nel contesto dominato dall'intelligenza artificiale, a causa dell'imprevedibilità del suo sviluppo, dell'impossibilità di coglierne i limiti e dell'attuale stato iniziale della sua regolamentazione — i cui effetti dovranno essere valutati solo dopo un congruo periodo di applicazione. Sembra pertanto logico ritenere che i minori siano, in misura ancora maggiore, soggetti a vulnerabilità.

Così, la vulnerabilità – intesa come debolezza, fragilità o anche insicurezza – che può riguardare gli adulti, si amplifica ulteriormente quando gli utenti o i destinatari sono minori; potremmo definirla, in tal caso, una vulnerabilità qualificata speciale. Si tratta di soggetti la cui inesperienza e immaturità li colloca in una posizione di maggiore debolezza, con strumenti più limitati per affrontare situazioni dannose e una scarsa capacità di percepire i pericoli o, eventualmente, le conseguenze delle proprie azioni. I rischi, pur astrattamente equivalenti a quelli corsi dagli adulti, crescono esponenzialmente a fronte delle maggiori difficoltà nel contrastarli e nel porvi rimedio. La particolare vulnerabilità dei minori nei contesti tecnologici – analogici o digitali – si intensifica ulteriormente nel mondo che inizia ora ad articolarsi intorno all'intelligenza artificiale, la quale, come vedremo più avanti, ha ormai invaso anche l'ambiente dei social network.

3. I minori e l'Intelligenza Artificiale: rischi specifici

Come già accennato, le diverse tecnologie basate sull'intelligenza artificiale possono comportare rischi significativi anche per gli adulti. Tuttavia, il presente studio si concentra in modo specifico sui pericoli che tali tecnologie rappresentano per i minori, nella consapevolezza che, in un'epoca caratterizzata da continue perturbazioni digitali, essi risultano particolarmente esposti a causa della loro – al momento – naturale immaturità e inesperienza. Per tale ragione, in queste pagine non verranno affrontate altre sfide, comprese quelle che riguardano la democrazia (come l'impatto sulla formazione dell'opinione pubblica) o quelle tipicamente riconducibili all'intervento degli adulti, come il processo decisionale automatizzato[4] in ambiti quali l'occupazione o i servizi bancari, dai quali i minori restano esclusi[5].

D'altra parte, è molto difficile immaginare tutti i contesti in cui l'intelligenza artificiale può

[4] Cfr. A. Papa, *L'utilizzo dell'intelligenza artificiale nell'adozione delle decisioni pubbliche: tecnica, politica e tutela dei diritti,* in *lce online* 2/2023.

[5] Cfr. F. Balaguer Callejón, *La costituzione dell'algoritmo e le trasformazioni culturali nella società digitale,* in *lce online* 3/2024.

costituire un pericolo per i bambini. Inoltre, chi non ha familiarità con queste tecnologie o non è consapevole della loro presenza in strumenti e oggetti sempre più comuni nella vita quotidiana, difficilmente sarà in grado di intuire che anche i minori possano esserne utenti o destinatari. In altre parole, chi non possiede almeno una conoscenza di base dei contesti in cui l'intelligenza artificiale è già integrata, tenderà a considerarla una tecnologia estranea e distante dal mondo dell'infanzia. Nulla di più errato, poiché dispositivi come *smartphone, smartwatch*, tablet o giocattoli intelligenti sono ormai largamente diffusi anche tra i bambini più piccoli, indipendentemente dal potere d'acquisto, dalla classe sociale o dal livello di istruzione della famiglia.

Com'è noto, lo strumento principale di cui disponiamo per affrontare questi rischi è la tutela della privacy[6], incarnata direttamente dal diritto alla protezione dei dati personali. Senza dubbio, la sfida più complessa si presenta proprio nell'utilizzo delle nuove tecnologie. I minori, inoltre, rappresentano un *target* di particolare interesse in un contesto commerciale come quello attuale, in cui la monetizzazione dei dati personali costituisce uno degli elementi chiave dei profitti[7] delle grandi aziende tecnologiche[8].

In questo contesto, la protezione dei dati rappresenta l'asse centrale attorno al quale ruota la tutela degli altri diritti. In altre parole, affrontare le minacce alla privacy è diventato il baluardo su cui si regge l'intera architettura della protezione delle persone negli ambienti digitali, compresi quelli fondati sull'intelligenza artificiale – quasi come se non vi fossero altri rischi il cui impatto sulla dignità della persona fosse, come vedremo nei paragrafi seguenti, ancora più diretto e tangibile.

3.1. Il mondo dei giocattoli intelligenti

Approfondendo la questione, il primo ambito a cui occorre rivolgere l'attenzione è quello dei giocattoli intelligenti (noti anche come *connected toys, smart toys* o *IoToys*). Si tratta di dispositivi che possono offrire diverse funzionalità: connettersi alla rete Wi-Fi o via *Bluetooth*, essere utilizzati tramite applicazioni per *smartphone* e interagire con l'utente attraverso tali risorse in modo sempre più coinvolgente, man mano che l'intelligenza artificiale viene implementata in questi prodotti[9]. Non è quindi necessario trovarsi di fronte a un dispositivo digitale in senso stretto – come un tablet, uno smartphone o una console – per riscontrare l'evoluzione tecnologica: anche giocattoli tradizionali come peluche o bambole si stanno trasformando in strumenti connessi e intelligenti[10].

---

[6] Lo si può osservare, ad esempio, in M . Arenas Ramiro e D. Díaz Lima, *Privacidad y derechos digitales*, in F. Balaguer Callejón e L. Cotino Hueso (Coords.), *Derecho público de la inteligencia artificial,* Zaragoza, 2023, p. 295 ss.: "De ahí que la privacidad se configure como uno de los pilares del Estado democrático. A mayor control personal del uso que se hace de sus datos personales, de su información personal, mayor participación ciudadana [...]. La privacidad, como ha quedado dicho, tiene un carácter transversal a los citados derechos digitales, ya sea porque implica un tratamiento de datos personales, ya sea porque la privacidad, la facultad de control sobre la propia información personal, actúa como instituto de garantía de los mismos".

[7] Afferma, L. Gómez Abeja, *Inteligencia artificial y derechos fundamentales*, cit., p. 94, che: "[c]omo es obvio, las posibilidades de control de esa información por parte de sus titulares se difuminan en un universo de trasiego constante y masivo de datos, por los que todos pujan. Y es que ello sucede sobre todo por el gran interés que existe en los datos, de los que se ha hablado como el petróleo del siglo XXI, pues se han convertido en un activo patrimonial de un valor económico sin precedente en los mercados".

[8] T. De la Quadra Salcedo, *Retos, riesgos y oportunidades de la sociedad digital*, in T. De la Quadra Salcedo, e J. L. Piñar Mañas (Dirs.), *Sociedad digital y Derecho,* Madrid, 2018, p. 44: "El impacto de las tecnologías digitales en el mundo económico es asombroso. No hay más que ver cómo en la actualidad en el *ranking* de las 10 mayores empresas del mundo por capitalización bursátil, 7 son tecnológicas, lo que supone que en sólo diez años se ha producido un cambio radical al sustituir a las empresas de banca y energía que ocupaban dichos puestos, por otras completamente distintas".

[9] Si veda a questo proposito R. Brito, P. Dias, e G. Oliveira, *Young children, digital media and smart toys: how perceptions shape adoption and domestication*, in *BJET,* 49/2018.

[10] Si rinvia, per un approfondimento, a O. P. Albuquerque, M. Fantinato, J. Kelner e A. P. De Albuquerque, *Privacy in smart toys: risks and proposed solutions*, in *Electronic Commerce Research and Applications,* 39/2020, p. 3: "The term smart toy covers a range of play products that can present different levels of network, processing and reasoning cap abilities. Smart toys appear in various shapes, including anthropomorphic and humanoid, such as a plush toy, a doll, a companion robot or

Come si può osservare, i potenziali utenti degli *smart toys* non sono gli adolescenti, sui quali generalmente si concentra l'attenzione quando si analizza l'impatto delle nuove tecnologie sulla vita dei minori – ragazzi e ragazze dai 12 anni in su, utenti abituali dei social network, perfettamente in grado di orientarsi in tali contesti[11] (spesso persino meglio dei propri genitori) e, in linea di massima, consapevoli dei rischi impliciti nel web 2.0, poiché queste applicazioni fanno ormai parte integrante della loro quotidianità. In questo caso, invece, ci troviamo di fronte a bambini in tenerissima età, che nascono completamente immersi nella tecnologia, figli di genitori che, a loro volta, iniziano a essere nativi digitali, e che già prima del compimento del primo anno di vita entrano in contatto con dispositivi come smartphone o tablet connessi a Internet[12].

Alla luce di questa circostanza, la dottrina è chiamata a elaborare strumenti capaci di garantire la tutela di tutti i diritti potenzialmente lesi, partendo dal presupposto che i genitori di questi piccoli utenti non sempre riescono a supplire alla naturale incapacità del bambino in tenera età (la maggior parte degli studi consultati si concentra sulla fascia 0-9 anni[13]) di intuire che l'interazione con un dispositivo apparentemente ludico può comportare rischi di cui, talvolta, nemmeno gli stessi adulti sono pienamente consapevoli. Non è nostra intenzione denunciare atteggiamenti negligenti da parte dei genitori – che ovviamente esistono[14] –, quanto piuttosto sottolineare che ci troviamo ancora una volta di fronte a una questione scarsamente approfondita dagli studiosi di diritto, rispetto alla quale resta da valutare se il GDPR e il RIA offrano una protezione davvero efficace.

La prima preoccupazione legata all'utilizzo dei giocattoli intelligenti da parte dei bambini riguarda la gestione dei loro dati personali. Questi dispositivi possono essere dotati di sensori, fotocamere, microfoni e localizzatori GPS, e sono quindi in grado di raccogliere, memorizzare, elaborare e condividere dati sensibili come immagini, voce e posizione. Nonostante l'implicazione di informazioni di tale rilevanza, la maggior parte della letteratura scientifica consultata concorda nel rilevare che i produttori non forniscono informazioni sufficienti ai genitori dei minori destinatari di questi giocattoli, né in merito alle modalità di trattamento dei dati raccolti, né rispetto alla coerenza del consenso prestato dagli adulti – compromessa non solo dalla scarsa trasparenza, ma anche dalla limitata (o assente) possibilità di configurare le impostazioni relative alla privacy dei dispositivi[15].

Tuttavia, va sottolineato che il rischio non riguarda esclusivamente la raccolta di dati con finalità di monetizzazione o di profilazione pubblicitaria mirata verso questo specifico gruppo di utenti –

a wearable gadget. Technological solutions used by these smart toys vary since Augmented Reality (AR) resources to advanced functions such as wireless connectivity features and AI-based conversation functions".

[11] Si veda, a questo proposito, Gómez Fernández, che offre un'interessante analisi della vulnerabilità dei minori, sottolineando come essi non manifestino particolari difficoltà nell'utilizzo delle nuove tecnologie: "Además, en el caso concreto que ahora nos ocupa, los niños, niñas y adolescentes no están en situación de desventaja respecto del resto de la sociedad en materia de uso de los entornos digitales, ni de dominio de las herramientas que esos entornos proporcionan. Se trata de nativos digitales que conocen y usan esos entornos con menos dificultad que los adultos mayores, y que comprenden e integran el lenguaje de las redes con mayor rapidez, adaptándose fácilmente a estos entornos", in I. Gómez Fernández, *Posición jurídica de la infancia y la adolescencia en los entornos digitales*, in F. Balaguer Callejón *et al.* (Dirs.), *Derechos fundamentales y democracia en el constitucionalismo digital,* Cizur Menor (Navarra), 2023.

[12] Cfr. L. Arnott, I. Palaiologou e G. Colette, *Digital devices, internet-enabled toys and digital games: the changing nature of young children's learning ecologies, experiences and pedagogies*, in *BJET*, 49/2018: "For example, from birth children have been found to observe screen-based communication such as video calls with distant grandparents [...] and before their first birthday can begin to autonomously manoeuvre the user interface by swiping [...] Toddlers are known to access photographs and videos; navigate Youtube, and interact with and complete some age appropriate digital games and applications [...] This research demonstrates the competence of very young children as users of digital media".

[13] Cfr. R. Brito, P. Dias e G. Oliveira, *Young children, digital media,* cit.; S. Chaudron *et al.* (eds.), *Kaleidoscope on the Internet of Toys – Safety, Security, privacy and societal insights - JCR Technical Reports,* Luxembourg, 2017.

[14] A questo proposito, commentano R. Brito, P. Dias e G. Oliveira, *Young children, digital media,* cit., che i bambini "explore portable devices mostly alone, while parents are occupied with work or house chores" (p. 814).

[15] Cfr., per tutti, O. P. Albuquerque, M. Fantinato, J. Kelner e A. P. De Albuquerque, *Privacy in smart toys,* cit.., p. 8: "The three most frequently found risks refer to: (1) the way in which the data is used (including its retention and mainly its disclosure); (2) the need for consent for any type of data operation to be performed (including offering different options); and (3) the need for broad transparency by smart toy makers so that the child guardian has an extensive knowledge of what is involved in the operations with the child's data".

conseguenze che, di per sé, meritano la massima attenzione al fine di individuare soluzioni efficaci –, ma comprende anche ulteriori pericoli legati alla sicurezza del minore e dell'intero nucleo familiare. Valente e Cárdenas, infatti, non si limitano ad analizzare i problemi in termini di privacy in senso stretto, ma pongono l'accento anche sulla vulnerabilità dei canali di trasmissione verso il cloud dei dati raccolti dal dispositivo. Tale debolezza può dar luogo a quello che gli autori definiscono *audio injection attack*, ovvero la possibilità che un attaccante introduca comandi vocali nel sistema, inducendo il giocattolo a richiedere ulteriori dati personali al bambino o a sollecitare azioni pericolose (come, ad esempio, l'apertura della porta di casa), con evidenti implicazioni in termini di rischio e sicurezza[16].

Per concludere l'analisi dei rischi connessi all'utilizzo degli smart toys, è opportuno segnalare un ulteriore aspetto critico. È evidente che una delle principali aspirazioni degli sviluppatori di tecnologie basate sull'intelligenza artificiale sia quella di rendere le applicazioni e le funzionalità il più possibile simili all'interazione umana. L'obiettivo, in tal senso, è che il dispositivo sia in grado di interagire con l'utente in modo analogo a quanto farebbe un altro essere umano. Ed è proprio questa la finalità valutata dal test di Turing: verificare la capacità di una macchina di esibire un comportamento – tramite domande e risposte – che risulti indistinguibile da quello di una persona. I progressi in questa direzione sono significativi, come dimostrano applicazioni già largamente conosciute, tra cui ChatGPT.

Parallelamente, le aziende stanno sviluppando robot umanoidi dal design sempre più realistico. Due elementi – da un lato, la capacità di risposta affinata grazie all'empatia e al riconoscimento delle emozioni attraverso l'analisi del linguaggio verbale e non verbale dell'utente; dall'altro, l'aspetto antropomorfo – rappresentano, secondo alcuni autori, una combinazione potenzialmente utile, ma che comporta anche rischi significativi, soprattutto per i bambini. Tra i possibili benefici, vari studi evidenziano come tali dispositivi possano contribuire allo sviluppo di competenze linguistiche, matematiche o sociali, fungendo da strumenti di supporto sia nell'ambiente scolastico che in quello domestico. A questi contributi positivi si contrappone però chi richiama l'attenzione non solo sui pericoli connessi al furto di dati o agli attacchi audio (molestie, adescamento, pedopornografia o rapimenti[17]), ma anche su problematiche relative alla socializzazione[18], alla sedentarietà e alla dipendenza[19]. Non va dimenticato che i destinatari principali sono bambini molto piccoli, i quali possono incontrare difficoltà nel distinguere l'umano dall'artificiale e che, sviluppando parte delle loro abilità sociali attraverso l'interazione con questi robot, potrebbero manifestare in futuro criticità, ad esempio nella risoluzione dei conflitti[20].

---

[16] Cfr. J. Valente e A. A. Cárdenas, *Security & Privacy in Smart Toys*, in *Proceedings of the 2017 Workshop on Internet of Things Security and Privacy,* p. 19: "In particular, we point out that while most of the concerns about smart toys have focused on privacy, our audio injection attack can potentially be more dangerous as it is targeting young children who are more vulnerable to deception and who presumably trust the smart toy. For example, the attacker can inject audio to the device so the smart toy tells children to open the door to their houses, or to change combination locks, or lies about their parents. The attacker can even be mean to the child, insulting their appearance or intelligence and therefore eroding from an early age their self-esteem and their trust of the toy and technology".

[17] Cfr. O. P. Albuquerque, M. Fantinato, J. Kelner e A. P. De Albuquerque, *Privacy in smart toys,* cit.., p. 12.: "Many examples of physical and psychological safety risks have been listed by researchers in a list that appears to be far from extensive. The creativity of criminals is certainly greater. Concerns with harassment, stalking, grooming, sexual abuse, exploitation, kidnapping, pedophilia and other types of violence were raised".

[18] R. Brito, P. Dias e G. Oliveira, *Young children, digital media,* cit., p. 816.: "For some parents, the use of digital media is an impediment for children to engage in other activi- ties, such as playing outdoors or socializing with family and friends and may also replace physical activity. Most parents believe in a balance of all activities — watching TV, playing digital games (of various types), running outdoors or playing with physical toys — as negative outcomes may result from excesses [...] The negative perceptions mainly focus on addiction and anti-socialization, in which children, when playing, alienate themselves from family members, not paying attention to them and these situations are sometimes the subject of family discussions".

[19] M. I. De la Iglesia Monje, *La inteligencia artificial aplicada a la robótica en los juguetes: el colectivo vulnerable de los menores*, in *Revista Crítica de Derecho Inmobiliario,* 788/2021, afferma che: "hay que valorar y cuantificar el hecho de que la tecnología influya en hacer menos humanos y más pasivos a los menores, influya en su autonomía personal, cree una adicción a estos juguetes y que a su vez provoquen una pérdida de su capacidad de iniciativa y creatividad".

[20] Pertanto M. C. Núñez Zorrilla, *Menor de edad e inteligencia artificial avanzada. Nuevos desafíos jurídicos*, in *Actualidad civil,* 12/2019, ritiene che: "la comunicación con el robot puede llegar a recrear un sistema de representación simbólica en

3.2. Algoritmi che raccomandano contenuti sui social network

Un ulteriore impiego dell'intelligenza artificiale che comporta rischi rilevanti per i minori e gli adolescenti riguarda i "sistemi di raccomandazione" – in grado di filtrare mediante utilizzo di algoritmi informazioni e fornire contenuti personalizzati per gli utenti – integrati in diversi social network, come Instagram o Facebook. Per comprendere appieno la portata del problema è necessario partire da una premessa fondamentale: gli adolescenti del XXI secolo hanno talmente interiorizzato i social network nella loro quotidianità da costruirvi una parte essenziale della propria identità e personalità. I nativi digitali di oggi sanno – in misura più o meno consapevole – a quali rischi sono esposti utilizzando queste applicazioni; tuttavia, l'importanza che attribuiscono a tali spazi virtuali li porta ad accettare tali rischi come parte del gioco, continuando a partecipare attivamente al contesto digitale. In effetti, le relazioni analogiche si estendono al mondo virtuale, i legami interpersonali si rafforzano online e si può affermare, quasi senza timore di smentita, che i membri di un gruppo che non sono presenti sulle piattaforme social incontrano maggiori difficoltà a partecipare alle dinamiche relazionali o a integrarsi pienamente in esse.

Partendo da questa premessa, va sottolineato che – secondo numerosi studi – l'età di accesso ai social network è sempre più precoce, con bambini che iniziano a utilizzarli già a partire dagli 8 anni[21]. In questa fase della vita, il processo di formazione della personalità si trova in un momento cruciale: gusti, interessi e hobby cominciano a delinearsi; gradualmente i minori maturano, costruendo un'immagine di sé e sviluppando un primo bagaglio di opinioni che riflette una visione iniziale del mondo, della società e, in ultima analisi, della propria esistenza.

E, naturalmente, in tutto questo processo di maturazione e socializzazione, l'ambiente digitale assume un ruolo sempre più centrale. I social network si configurano come spazi in cui condividere preoccupazioni, interagire con i coetanei, stringere nuove relazioni ed esplorare interessi emergenti. Ed è proprio su questi ultimi che si concentra l'attenzione dei sistemi di raccomandazione integrati nelle piattaforme social. Gli utenti manifestano le proprie preferenze verso determinati contenuti sia attraverso azioni volontarie e consapevoli – ad esempio, mettendo "mi piace" a un post, condividendolo, salvandolo in una sezione del proprio profilo o seguendo l'autore –, sia mediante comportamenti involontari, come il semplice soffermarsi per qualche secondo su un contenuto. L'intelligenza artificiale, tramite algoritmi di apprendimento automatico (*machine learning*[22]), registra, elabora e restituisce tali informazioni sotto forma di suggerimenti personalizzati, noti come *feed*[23].

---

el que se halla implícita una determinada concepción del mundo y se despliega todo un sistema de valores... la IA puede convertirse en una escuela de emociones y de comportamientos para los niños, proyectando sobre estos una imagen idealizada del mundo y de sí mismos".

[21] È quanto emerge dal rapporto *The Common Sense Census: Media Use by Tweens and Teens 2021,* pubblicato nel marzo 2022, sebbene sia opportuno precisare che i dati si riferiscono agli Stati Uniti (https://www.commonsensemedia.org/sites/default/files/research/report/8-18-census-integrated-report-final-web_0.pdf). In Spagna, l'età media di accesso a un dispositivo mobile personale con connessione a Internet è di 10,96 anni. Si veda, a questo proposito, M. Morillas Fernández, *La protección de datos, atendiendo a la edad de los usuarios, como garantía en la prevención de los riesgos de la sociedad digital*, in *Revista Internacional de Doctrina y Jurisprudencia,* 29/2023.

[22] Concordiamo con quanto affermato da A. C. Bueno Fernandes e E. Campos Pellanda, *Estereotipos de género en TikTok e Instagram: un experimento de ingeniería inversa para entender los mecanismos de los algoritmos de las redes sociales*, in *Universitas-XXI: Revista de Ciencias Sociales y Humanas de la Universidad Politécnica Salesiana del Ecuador,* 37/2023, secondo cui comprendere e discutere gli algoritmi che plasmano sempre più il nostro modo di vedere il mondo e il nostro senso della realtà è fondamentale per comprendere gli impatti negativi delle nuove tecnologie e come mitigarli.

[23] K. D. Calva Cabrera, T. B. León Alberca e C. G. Arpi Fernández, *Capítulo 1. Inteligencia Artificial en las redes sociales digitales*, in *Espejo de Monografías de Comunicación Social,* 23/2024, sostengono che, nonostante ciò: "se debe recordar que los *feeds* de noticias dentro de las redes sociales digitales se personalizan de acuerdo al comportamiento del usuario – este trabajo es una responsabilidad de la IA, pues su labor es de relevancia – y en este entorno se convierten en el conjunto de algoritmos de aprendizaje automático que permiten predecir el comportamiento o gustos de un usuario, es decir [...] se orienta a proporcionar y atender los requerimientos puntuales", in. In tal senso si esprimono anche C. Rodríguez Gordo e J. Sarrión Andaluz, *Polarización social y control político: algunas consecuencias de la inteligencia artificial y las redes sociales para la razón práctica*, in *Cuadernos salmantinos de filosofía,* 48/2021, i quali sostengono che: "los algoritmos que procesan

È a questo punto che emergono alcune questioni critiche. La prima riguarda l'età minima necessaria affinché un minore possa esprimere validamente il consenso al trattamento dei propri dati personali da parte dei fornitori di servizi digitali. Come noto, si tratta di un tema ricorrente, già affrontato dal GDPR, il cui articolo 8.1 stabilisce che il consenso è valido solo se espresso da soggetti di almeno 16 anni, salvo che gli Stati membri decidano di abbassare tale soglia fino a un minimo di 13 anni. In Spagna, ad esempio, la Ley Orgánica 3/2018, del 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD) ha fissato l'età a 14 anni, anche se l'attuale governo ha manifestato l'intenzione di elevarla nuovamente a 16 anni, come emerge dal "Anteproyecto de Ley Orgánica para la protección de las personas menores de edad en los entornos digitales"[24]. In ogni caso, l'efficacia concreta di queste disposizioni è discutibile, poiché i fornitori di servizi raramente adottano misure efficaci per verificare l'età reale degli utenti.

La seconda questione problematica, strettamente connessa alla precedente, riguarda il contenuto stesso del *feed* di raccomandazioni nei social network. Se, come abbiamo visto, i minori accedono a questi servizi in età sempre più precoce, dispongono di dispositivi personali e il monitoraggio da parte degli adulti risulta complesso – anche in ragione della necessità di strumenti come le applicazioni di parental control, che non tutti conoscono, che molti non considerano efficaci o desiderabili, e che, peraltro, possono entrare in tensione con altri diritti fondamentali, come la privacy del minore –, è plausibile ritenere che una parte di questi utenti possa essere esposta a contenuti non adeguati, nel complesso, al proprio livello di maturità.

Tra i contenuti che destano maggiore preoccupazione figura innanzitutto la pornografia[25]. Turienzo Fernández evidenzia che, secondo il rapporto di Save the Children Spagna *"(Des)información sexual: pornografía y adolescencia"* (p. 27), pubblicato nel 2020, quasi sette minori su dieci tra i 13 e i 17 anni (precisamente il 68,2%) consumano regolarmente contenuti pornografici, aggiungendo che nel nostro Paese l'età media del primo contatto con la pornografia si colloca tra gli 8 e i 12 anni[26]. Si tratta, dunque, di un'esposizione libera e non mediata a contenuti sessuali in una fase in cui identità e personalità sono in piena formazione, spesso in assenza di qualsiasi esperienza affettiva o sessuale reale, e senza alcun tipo di mediazione o orientamento adulto.

In tal senso, una prima conseguenza evidenziata è che *"il consumo precoce di contenuti pornografici ha un impatto sul tipo di atteggiamenti degli individui, soprattutto degli uomini nei confronti delle donne"*. Studi condotti su campioni rappresentativi di adolescenti tra i 14 e i 17 anni rivelano una correlazione tra la fruizione regolare di pornografia *online* e la perpetrazione di comportamenti coercitivi o abusivi, in particolare da parte dei ragazzi[27]. Non sorprende, dunque, che l'aumento dei casi di violenza sessuale tra minori e adolescenti abbia spinto la dottrina a interrogarsi se, e in quale misura, il consumo precoce di pornografia possa incidere su tale fenomeno[28].

A ciò si aggiunge una propensione a interiorizzare modelli sessuali distorti: pratiche prive di consenso, sesso non protetto, dinamiche di dominio e, in parallelo, sentimenti di insoddisfazione rispetto alla propria immagine corporea o al proprio ruolo nei rapporti intimi. In sintesi, un problema che fino a poco tempo fa era prevalentemente circoscritto a siti web specializzati – spesso già privi di reali strumenti di limitazione dell'accesso – si è ormai esteso ai social network, che hanno attraversato

---

la información obtenida de las redes sociales se encargan de decidir qué publicaciones visualiza un usuario en su móvil con relación a su actividad y realizar un análisis acerca de su comportamiento para poder optimizar el primer proceso".

[24] https://www.mpr.gob.es/servicios/participacion/Documents/ANTEPROYECTO%20DE%20LEY%20ORGÁNICA%20PARA%20LA%20PROTECCIÓN%20DE%20LAS%20PERSONAS%20MENORES%20DE%20EDAD%20EN%20LOS%20ENTORNOS%20DIGITALES.pdf.

[25] Uno studio recente e molto completo lo troviamo in A. Turienzo Fernández, *Pornografía y daño a menores: principales efectos perjudiciales asociados a la exposición temprana a pornografía*, in *Indret: Revista para el análisis del Derecho,* 2/2024.

[26] Secondo la campagna "Por no, porno. Educa a tus hijos antes de que el porno lo haga por ti", della Fundación de Ayuda contra la Drogadicción (https://fad.es/campana/pornoeducar/).

[27] C. Muñoz Villanueva, *Consumo de pornografía y normalización de conductas violentas en las relaciones sexuales de los jóvenes*, in *Atlánticas – Revista Internacional de Estudios Feministas,* 9(1)/2024.

[28] A. Turienzo Fernández, *Pornografía y daño a menores*, cit., p. 453.

un evidente processo di "pornificazione". Questo fenomeno si manifesta attraverso l'elusione delle condizioni d'uso, con piattaforme come Instagram, TikTok e soprattutto X (ex Twitter), che si distingue per la quasi totale assenza di filtri o limitazioni su questo tipo di contenuti[29].

Tuttavia, i contenuti inappropriati per i minori non si limitano alla sfera pornografica. Una seconda categoria comprende le pubblicazioni che minacciano la convivenza civile e i valori fondamentali della democrazia: messaggi d'odio, discriminatori, violenti. Sui *social network* proliferano profili che promuovono ideologie apertamente antidemocratiche e radicali, che, rifacendosi al diritto alla libertà di espressione[30], diffondono contenuti xenofobi, aporofobici, misogini o ostili nei confronti della comunità LGTBI.

È opportuno ricordare, a questo punto, che gli algoritmi delle piattaforme social selezionano e raccomandano i contenuti sulla base dell'interesse mostrato dagli utenti e, inoltre, privilegiano quelli che generano maggiore coinvolgimento. Le pubblicazioni controverse, infatti, tendono a suscitare più interazioni – "mi piace", commenti, condivisioni – e, di conseguenza, vengono premiate in termini di visibilità[31]. In tal modo, contenuti polarizzanti e pericolosi finiscono per essere amplificati, contribuendo a una normalizzazione di discorsi e atteggiamenti contrari ai principi costituzionali di uguaglianza, dignità e non discriminazione.

Come primo esempio, tra i numerosi disponibili, si può citare l'interessante studio condotto da Gil Ramírez e Gómez de Travesedo Rojas, che analizza il trattamento riservato ai minori stranieri non accompagnati nei contenuti pubblicati su YouTube. L'indagine evidenzia la presenza di video che diffondono opinioni soggettive prive di verifica, basate su stereotipi alimentati da posizioni ideologiche radicali. In essi si fa riferimento, ad esempio, a un presunto legame con la delinquenza, a pratiche religiose fanatiche o a valori considerati incompatibili con la democrazia. Si tratta di contenuti liberamente accessibili, privi di qualsiasi filtro o limitazione, che veicolano una narrazione non pluralistica, lesiva della dignità delle persone coinvolte e, in definitiva, in contrasto con i principi fondamentali dello Stato democratico[32].

Un altro esempio significativo è rappresentato dalla normalizzazione di atteggiamenti maschilisti

---

[29] Cfr. B. Sánchez Rubio, *Pornografía, redes sociales y violencia sexual*, in C. Torres Fernández, J. Puig Guisado, G. Molano Caro, e A. I. Aguilar Cuesta (Coords.), *Diversidad sexual y género a través de la educación y las artes,* Madrid, 2021.

[30] Particolarmente rilevante appare la riflessione proposta da J. J. Marabel Matos, *Delitos de odio y redes sociales: el Derecho frente al reto de las nuevas tecnologías*, in *Revista de Derecho UNED,* 27/2021, secondo la quale "[e]l ordenamiento constitucional protege incluso a quienes lo niegan, ya que en términos generales el derecho a la libertad de expresión admite la protección de cualquier idea, por rechazable, perniciosa o aberrante que objetivamente pudiera parecer a muchos desde la perspectiva de los valores éticos, pues así lo requieren el pluralismo, la tolerancia y el espíritu de apertura, sin los cuales no existiría la sociedad democrática. Ahora bien, esta misma doctrina constitucional establece que la libertad de expresión no es un derecho absoluto, sino que está sujeta a las restricciones que pongan en manifiesto peligro la convivencia pacífica de los ciudadanos o la protección de otros bienes jurídicos, como la libertad y la dignidad de las personas".

[31] Cfr. L. Sell Trujillo, T. Núñez Domíngueze C. Romo Parra, *La violencia que seremos: jóvenes y taxonomía de mensajes de odio en las redes*, in *Espejo de Monografías de Comunicación Social,* 16/2023, p. 93: "Las redes sociales, por su propia inercia e intencionalidad explícita de conseguir visibilidad, tienen un elemento perverso relacionado con cómo se 'premian' los comentarios más arriesgados o generadores de controversia [...] La polémica genera visibilidad, que se traduce en el aumento de 'visualizaciones' o 'seguidores'. A mayor impacto, mayor relevancia independientemente de que el impacto se haya generado por contenidos positivos o negativos/agresivos (que suelen ser más capitalizables)".

[32] Le autrici sottolineano che la "red social YouTube estará desempeñando, en este sentido, una función de difusor de un discurso de odio [...] hacia la figura de los MENA que llegan o están en territorio español. El discurso de odio implica el uso de un lenguaje que considera al migrante como un sujeto antitético a los valores nucleares nacionales y como una amenaza a los valores que cimientan la convivencia social, y promueve mensajes que alimentan el rechazo, el menosprecio, la humillación, el descrédito y la estigmatización del individuo o colectivo en base a atributos que van desde la nacionalidad hasta la orientación sexual [...] En el caso de YouTube este discurso de odio se presenta construido principalmente a través de una narrativa emocional basada en opiniones subjetivas de ciudadanos particulares y/o anónimos, que no recurren a fuentes autorizadas para validar sus argumentos. [...] detectaron igualmente el uso de apreciaciones subjetivadas carentes o faltas de pluralidad de fuentes en los discursos de usuarios o colaboradores de medios convencionales hacia los inmigrantes", in M. Gil Ramírez e R. Gómez De Travesedo Rojas, *Estrategia discursiva sobre los MENA en YouTube. Construcción de un discurso de odio*, in *Revista Latina de Comunicación Social,* 80/2022.

all'interno delle relazioni di coppia tra adolescenti[33]. Sebbene sia noto che la prevenzione della violenza di genere è ormai presente nei programmi scolastici, contrastare l'influenza dei social network si rivela estremamente complesso. In questi ambienti digitali, il flusso continuo di contenuti che minimizzano – o persino negano apertamente – la gravità di questo fenomeno rappresenta un ostacolo rilevante. Va inoltre sottolineato che tali messaggi non vengono sempre veicolati in forma diretta o aggressiva, ma spesso si diffondono attraverso strumenti più sottili, come l'umorismo, una dinamica che, proprio per la sua apparente innocuità, può risultare ancora più pericolosa in quanto rafforza stereotipi e pregiudizi in modo subdolo e difficilmente percepibile dai destinatari[34].

Infine, è opportuno soffermarsi su un'ultima categoria di contenuti potenzialmente dannosi per i minori che utilizzano i social network: quelli relativi alla salute, sia fisica che mentale. Piattaforme come Instagram, Facebook o Snapchat promuovono profili e pubblicazioni di persone considerate di successo, che spesso portano all'estremo la cura del corpo e diffondono modelli di bellezza irrealistici. Bambini e adolescenti, in particolare le ragazze, accedono con facilità a tali contenuti che, pur apparendo innocui – poiché non esplicitamente violenti, discriminatori o sessuali – possono suggerire abitudini[35] finalizzate al miglioramento dell'aspetto fisico, comprese pratiche alimentari. Tali raccomandazioni, se non basate su indicazioni fornite da professionisti della salute e dell'alimentazione, attraverso una valutazione personalizzata delle condizioni del soggetto, possono rappresentare un rischio concreto per il benessere psicofisico dei minori.

Per quanto riguarda la salute mentale, è facilmente intuibile come uno dei principali rischi sia rappresentato dalla dipendenza. Prima di affrontare direttamente questo aspetto, è utile ricordare che l'algoritmo propone contenuti in base agli interessi dimostrati dall'utente. Questo significa che anche situazioni negative o patologiche possono essere rafforzate e amplificate. Si pensi, ad esempio, al caso di un adolescente con una storia di disturbi alimentari o con una condizione depressiva che comporta tendenze autolesive o pensieri suicidari. Se questo utente ricerca e interagisce con contenuti dannosi – come strategie per evitare di mangiare, perdere peso rapidamente, autolesionarsi senza lasciare segni o persino metodi per togliersi la vita[36] – l'algoritmo memorizza e rielabora queste informazioni,

---

[33] Il maschilismo strutturale, in generale, e la violenza di genere, in particolare, costituiscono una questione di grande rilevanza nell'attualità, con una crescente preoccupazione per la riproduzione di modelli tra i e le più giovani. Particolarmente interessante, in linea con quanto trattato nel presente lavoro, è lo studio di A. M. Pérez Vallejo, *Violencia de género digital en adolescentes. Previsiones legislativas y estrategias europeas para su abordaje y prevención*, in G. Tomás Martínez e A. Vidu Afloarei (Coords.), *Mujer como motor de innovación jurídica y social,* Valencia, 2021. L'autrice conclude (p. 618) che "no cabe duda que las conductas de dominio y sumisión entre parejas adolescentes en redes sociales y telefonía móvil están menos exploradas y requieren una mayor investigación. Es preciso realizar estudios que arrojen nuevas evidencias y cuestionen la influencia de la tecnología digital y de los medios de comunicación audiovisuales en el mantenimiento de estereotipos de género y patrones violentos sexualizados en línea; y el impacto que puedan tener para propiciar o favorecer la violencia de género en el entorno digital".

[34] Pertanto S. Civila de Dios *et al.*, *TikTok y la caricaturización de la violencia en las relaciones de parejas adolescentes*, in *Anàlisi: Quaderns de comunicació i cultura,* 69/2023, sostengono che: "este estudio arroja luz sobre la forma en que TikTok utiliza el humor y la parodia para representar la VNA *[violencia en el noviazgo adolescente]*, lo que conlleva riesgos de refuerzo de estereotipos, normalización de la violencia, promoción de relaciones tóxicas y trivialización del problema. Además, se subraya la influencia de TikTok en la percepción de la realidad y la dificultad de identificar la VNA, debido a su representación minimizada en la plataforma".

[35] Cfr. T. De Lima Burgos, L. Sousa Di Lauro e M. Diógenes Moreira, *Instagram e a reafirmação dos estereótipos de corpo ideal feminino*, in *Razón y palabra,* 24(109)/2020: "em um cenário em que as mulheres compõem o número predominante de usuários na plataforma do Instagram e que, historicamente, sofrem coerções estéticas imperativas e latentes, o público feminino acaba incorporando novas concepções de beleza, influenciadas pelo que é disseminado na mídia, especialmente nas redes sociais. No caso relatado neste estudo, a simples busca por três hashtags relacionadas a esses protótipos de beleza corporal foi o bastante para ocupar o campo visual do usuário com postagens relacionadas à temática, exibindo freneticamente imagens de mulheres magras, saradas e aparentemente sem marcas na pele, como estrias, celulites, ou quaisquer outros rastros de excessos de gordura, além de receitas e soluções, como dietas e exercícios físicos, para adquirir a silhueta considerada como bela, saudável e perfeita para o corpo feminino".

[36] È vero che, quando questo tipo di comportamenti ha iniziato a diffondersi, alcuni social network, come Instagram o, all'epoca, anche Twitter, hanno implementato sistemi di protezione, principalmente tramite link a pagine di assistenza. Tuttavia, esiste ancora un ampio margine di miglioramento, ad esempio nel rendere più rapidi i processi di rimozione dei contenuti.

generando nuove raccomandazioni coerenti con tali interessi. Si tratta di un meccanismo che, pur non essendo intenzionale, può aggravare condizioni cliniche già fragili, esponendo i minori a un rischio elevatissimo, per il quale è urgente individuare ed attuare misure efficaci di prevenzione e protezione.

Tuttavia, oltre al problema appena descritto – che si colloca in continuità con quelli già analizzati –, l'attività svolta sui social network esercita un'elevata attrattiva sugli utenti, in particolare sui minori, che vi trovano molteplici modalità di intrattenimento. In relazione al sistema di raccomandazione algoritmica, tale attrazione si intensifica per due ragioni principali: da un lato, come già osservato, il feed si costruisce in base alle preferenze individuali, elaborate a partire dal comportamento digitale dell'utente; dall'altro, la presentazione dei contenuti è progettata per massimizzare il tempo di permanenza sulla piattaforma. A tal fine, il formato proposto è altamente coinvolgente: brevi video con contenuti eterogenei – umoristici, parodici, sentimentali, sportivi, legati ai videogiochi o di carattere sessuale/ipersessualizzato. Varietà e velocità rappresentano due elementi centrali per una generazione abituata a sperimentare molto in tempi brevi, incline ad annoiarsi rapidamente e immersa in un ritmo di vita accelerato. È quindi frequente che l'accesso alla scheda delle raccomandazioni, inizialmente concepito come pausa di pochi minuti, si traduca in una permanenza prolungata, spesso inconsapevole.

In questo senso, la dipendenza da dispositivi digitali e social network – o, quantomeno, il loro utilizzo intensivo – è un fenomeno su cui esistono già numerosi studi, i quali evidenziano una serie di conseguenze dannose per gli utenti[37]. Il livello di dipendenza osservato è tale che alcune analisi equiparano i comportamenti degli individui dipendenti dalle nuove tecnologie a quelli delle persone affette da dipendenze da sostanze chimiche[38]. Tra le problematiche più frequentemente riscontrate si annoverano la perdita di autocontrollo, l'eccessiva preoccupazione per la connessione mobile, l'uso in contesti inappropriati o pubblici, disturbi dell'umore, conflitti interpersonali, la tendenza a privilegiare l'interazione con i dispositivi rispetto alle relazioni umane[39], il calo del rendimento scolastico, nonché sintomi quali ansia, depressione o panico sociale[40].

In conclusione, abbiamo esaminato i diversi tipi di contenuti ai quali i minori possono accedere con maggiore o minore facilità, contenuti che l'algoritmo memorizza, analizza ed elabora sulla base dell'interazione dell'utente con le pubblicazioni o i profili proposti, con l'obiettivo ultimo di incrementare il coinvolgimento (*engagement*). In questo quadro emergono due problematiche principali. La prima riguarda il consenso: è evidente che un numero significativo di minori utilizza quotidianamente i *social*

---

[37] M. I. Isidro De Pedro e T. Moreno Martín, *Redes sociales y aplicaciones de móvil: uso, abuso y adicción*, in *IJODAEP,* 4(1)/2018, p. 211: "No podemos, ni pretendemos, etiquetar de abuso o adicción los comportamientos de los sujetos de la muestra en cuanto al uso que hacen de las Redes Sociales y el *Whatsapp.* Sin embargo, de acuerdo con los resultados obtenidos, sí podemos afirmar que su uso es bastante elevado, así como también es alta la frecuencia general con la que se utilizan. De este modo, estas herramientas se han convertido en una parte importante de la vida cotidiana de las personas teniendo, casi constantemente, el móvil encima y estando, en muchos casos, conectados siempre a las redes. Y, por ende, esto puede suponer una limitación importante en sus vidas, generando dependencia y obstaculizando, perjudicando o dificultando otros aspectos de su vida social".

[38] A. G. Tresáncoras, C. García Oliva e J. A. Piqueras, *Relación del uso problemático de whatsapp con la personalidad y la ansiedad en adolescentes*, in *Health and Addictions,* 17(1)/2017, p. 28: "Griffiths [...] habló de adicciones conductuales a través del Modelo de Adicción Biopsicosocial, desde el que compara a personas adictas a sustancias químicas (tabaco, alcohol y otras drogas) con personas que interactúan con nuevas tecnologías u otros comportamientos reforzantes, concluyendo que la sintomatología es similar en ambos casos Estos síntomas de dependencia se manifiestan en modo de aislamiento, bien estando en compañía o solo, bajo rendimiento de estudios o el trabajo y atención centrada en el uso de las tecnologías [...] También aparece pérdida de control, búsqueda desesperada por tener refuerzos rápidos y sentimientos de malestar si no usa la tecnología suficiente [...] En la misma línea, Young [...] expone que el uso desmedido o abusivo puede llevar a un uso patológico y de ahí derivar a una adicción".

[39] Cfr. l'interessante studio di C. Simó Sanz *et al., Instrumentos de evaluación del uso problemático del teléfono móvil/ smartphone*, in *Health and Addictions,* 17(1)/2017.

[40] J. Fleta Zaragozano e J. L Olivares López, *Adolescencia y adicción a las nuevas tecnologías*, in *Boletín de la Sociedad de Pediatría de Aragón, La Rioja y Soria,* 39(3)/2009, p. 86: "se han descrito casos con bajo rendimiento escolar, síntomas depresivos, ansiedad y pánico social y, en casos más graves, trastornos de la personalidad y de la conducta e incluso el denominado trastorno de déficit de atención e hiperactividad (TDAH)".

network in violazione delle disposizioni previste dal GDPR e dalla LOPDGDD, senza che i fornitori di servizi abbiano posto in essere misure efficaci per contrastare tale fenomeno. La seconda concerne la presunta neutralità algoritmica: come abbiamo visto, l'algoritmo non si limita a proporre ciò che l'utente ha già mostrato di preferire (indipendentemente dall'adeguatezza rispetto alla sua età), ma privilegia i contenuti suscettibili di generare maggiore interazione, in una logica puramente commerciale.

La realtà è che né i fornitori di servizi né i loro modelli di business sembrano contemplare – in modo concreto ed efficace – la possibilità di controllare l'accesso dei minori alle piattaforme, né, tanto meno, quella di sviluppare sistemi algoritmici neutri o educativi. Con ciò intendiamo una tecnologia che, pur elaborando i dati e gli interessi degli utenti minorenni, sia orientata a fornire raccomandazioni coerenti con il loro processo di formazione della personalità, proponendo contenuti rigorosi, imparziali e pluralisti, ispirati a valori democratici come l'uguaglianza e la non discriminazione. Una tale impostazione – ci sia concessa l'espressione – forse non risponde pienamente agli "interessi" commerciali dei minori, ma certamente è conforme al loro interesse superiore. Purtroppo, si tratta ancora di una visione utopica, giacché questo tipo di algoritmo, orientato alla tutela, non genera i profitti attesi dai modelli economici attuali.

3.3. Deepfakes

L'ultimo rischio al quale sono esposti i minori, sia in quanto utenti sia in quanto destinatari di applicazioni basate sull'intelligenza artificiale, è rappresentato dai deepfake. Secondo García Ull, un *deepfake* è "un video iperrealistico manipolato digitalmente per rappresentare persone che dicono o fanno cose che in realtà non sono mai accadute" [41]. Meneses, a partire da alcune immagini diventate virali in cui Nancy Pelosi appariva ubriaca, sottolinea la necessità di fornire una definizione canonica del concetto di *deepfake*, osservando che tali immagini costituiscono in realtà un esempio di *cheapfake* – ovvero un video rallentato digitalmente per creare questo effetto senza l'intervento dell'intelligenza artificiale. Dal suo punto di vista, è ancora necessario proporre una concettualizzazione chiara e precisa del termine, che dovrebbe includere necessariamente un riferimento all'uso dell'intelligenza artificiale e alla falsità del contenuto rappresentato[42].

La maggior parte degli studi consultati si concentra sul pericolo che l'impiego di questo tipo di intelligenza artificiale rappresenta per la democrazia. In effetti, è innegabile che il livello di perfezione raggiunto dai *deepfake* costituisca un rischio estremamente serio: si pensi, ad esempio, alla possibilità di creare un video del tutto realistico in cui una figura istituzionale (come un Capo di Stato, un Presidente del Governo, o un leader dell'opposizione) compia atti inappropriati o rilasci dichiarazioni sconvenienti su temi rilevanti dell'agenda politica. Tuttavia, in linea con quanto anticipato in apertura, non approfondiremo questo aspetto, poiché non si configura, in linea di principio, come un ambito direttamente riconducibile all'esperienza dei minori.

D'altra parte, il lettore potrà facilmente immaginare gli usi di questa tecnologia che comportano un rischio concreto per i minori. In contesti di cyberbullismo scolastico, ad esempio, essa può essere impiegata per far "pronunciare" alla vittima affermazioni che ne accentuino l'isolamento o ne favoriscano l'emarginazione dal gruppo, per attribuirle giudizi di valore che la allontanino dall'unico sostegno di cui dispone nella vita quotidiana (i compagni di classe e, talvolta, anche gli insegnanti), o semplicemente per costruire dichiarazioni suscettibili di diventare virali e screditarla pubblicamente. Questo uso dell'intelligenza artificiale, in combinazione con tecniche come il *morphing* e la creazione di caricature, costituisce un ulteriore strumento a disposizione dei bulli per esercitare violenza nei

[41] Cfr. F. J. García Ull, *Deepfakes: el próximo reto en la detección de noticias falsas*, in *Quaderns de comunicación i cultura,* 64/2021.

[42] J. P. Meneses, *Seeking to define deepfakes form U.S. state laws*, in *Comunicación y Sociedad,* 37(3)/2024, afferma a questo proposito: "It seems clear to us that any conceptualization that may be made at this point must include: i) a generic or detailed reference to AI; ii) the idea of manipulated or fake content; and iii) the four supports *[vídeo, sonido, imagen y texto]* or, alternatively, a generic and comprehensive formulation".

confronti delle loro vittime *online*, prolungandone la sofferenza oltre i confini spazio-temporali dell'ambiente scolastico.

In secondo luogo, e con una marcata dimensione di genere, si segnala la creazione di immagini a contenuto sessuale raffiguranti ragazze e giovani donne adolescenti, fenomeno noto come "abuso di immagini intime non consensuali". A partire da fotografie generalmente reperite online (ritraenti corpi nudi e/o impegnati in pratiche sessuali), l'intelligenza artificiale è in grado di sovrapporre in modo sempre più sofisticato il volto della vittima, ottenendo un risultato pressoché indistinguibile dalla realtà. Tale grado di verosimiglianza è accentuato dalla possibilità di simulare anche la voce della persona ritratta: quanto maggiore è la disponibilità di frammenti audio, tanto più accurata sarà la riproduzione da parte dell'intelligenza artificiale[43].

La variabile di genere sembra incidere in modo significativo, come evidenziato da diversi studi che sottolineano non solo l'entità numerica del fenomeno, ma anche il livello di perfezione raggiunto dall'applicazione. In altri termini, i programmi basati sull'intelligenza artificiale impiegati per l'elaborazione e la fusione di immagini producono risultati sensibilmente più verosimili quando i soggetti rappresentati sono donne[44], proprio perché l'efficacia dell'operazione tende a perfezionarsi quanto più frequentemente viene eseguita.

4. Conclusioni

In queste pagine si è cercato di delineare la posizione dei minori in un contesto in cui l'intelligenza artificiale è sempre più pervasiva, a partire dall'analisi di applicazioni concrete effettivamente accessibili e utilizzabili da parte loro.

Innanzitutto, occorre partire dal presupposto che i bambini e gli adolescenti del XXI secolo, in quanto nativi digitali, crescono circondati dalle nuove tecnologie, con le quali iniziano a interagire fin dalla più tenera età. Tuttavia, il mondo virtuale, incluso quello del Web 2.0, comporta una serie di rischi che non possono essere ignorati. Sebbene questo gruppo dimostri una notevole dimestichezza nell'utilizzo di applicazioni come Instagram, X o Facebook – spesso superiore a quella dei loro stessi genitori –, molti dei pericoli che li colpiscono nell'ambiente analogico trovano una continuità, oltre che nuove forme di manifestazione, anche in quello digitale. La particolare vulnerabilità dei minori si riflette pertanto nei social network, non tanto per una mancanza di competenze tecniche, quanto piuttosto per l'assenza delle risorse e degli strumenti necessari ad affrontare situazioni quali grooming, stalking, cyberbullismo o sexting, a causa della loro immaturità ancora in via di superamento.

Questa particolare vulnerabilità si accentua ulteriormente quando l'ambiente in cui i minori si muovono non corrisponde più alla fase tecnologica precedente, rappresentata dal Web 2.0, bensì a quella attuale, caratterizzata essenzialmente dall'applicazione e dallo sviluppo dell'intelligenza

---

[43] A questo proposito, C. Trujillo Cabrera, *El derecho a la propia imagen (y a la voz) frente a la inteligencia artificial*, in *InDret: Revista para el análisis del Derecho,* 1/ 2024, osserva che: "[p]oca duda cabe a este respecto cuando ese uso de la voz o imagen implica la imputación de hechos o la manifestación de juicios de valor a través de acciones o expresiones que de cualquier modo lesiones la dignidad de la persona, menoscabando su fama o atentando contra su propia estimación [...] Es lo que sucede, por ejemplo, en los casos de lo que se conoce como abuso de imagen íntima no consentida (*non-consensual intimate image abuse*), que, de acuerdo con el Instituto Europeo para la Igualdad de Género (conocido por sus siglas en inglés, EIGE, *European Institute for Gender Equality*) consiste en la creación no consentida y/o la difusión no consentida, principalmente en línea, de imágenes/vídeos íntimos o privados o de carácter sexual, que, en su gran mayoría, además, tienen a la mujer como objetivo. Estos supuestos están cada vez más presentes en la sociedad, afectando tanto a personas famosas como a personas anónimas".

[44] A. Vara López, *Los sesgos de género en la inteligencia artificial: deepfakes y deepnudes*, in C. Prados García (Dir.), *Igualdad de género en entornos digitales. Desafíos educativos*, Madrid, 2024, mette in guardia su "la existencia de aplicaciones como *DeepNude,* creadas con el propósito explícito de 'desnudar' a una mujer. Solo se necesita un retrato de la víctima y, a partir de una base de datos de 10.000 fotos de mujeres desnudas tomadas de internet, se le añade un cuerpo sin ropa. El resultado suele ser mucho más logrado cuando las víctimas son mujeres, debido a la abundancia de desnudos femeninos en la red".

artificiale. Viviamo oggi in un'epoca segnata dall'incertezza, in cui possiamo soltanto intuire – ma non confermare con certezza – fino a che punto si spingerà l'evoluzione dell'intelligenza artificiale, con teorie che spaziano da ipotesi di sviluppo limitato fino a scenari in cui essa superi ampiamente le capacità umane, raggiungendo un'autonomia propria. In tale contesto, anche gli adulti si trovano esposti a una pluralità di rischi (compresa la stessa tenuta del sistema democratico, sebbene quest'ultimo aspetto non sia stato oggetto di analisi in questo contributo), risultando vulnerabili rispetto a diverse applicazioni dell'intelligenza artificiale, come il machine learning, i big data o i processi decisionali automatizzati. È proprio l'intersezione tra questi due elementi – da un lato, la particolare vulnerabilità dei minori già osservata nella fase tecnologica precedente, e dall'altro, la vulnerabilità degli adulti di fronte all'attuale scenario – che ci autorizza a parlare di qualcosa che va oltre la "vulnerabilità speciale" o la "vulnerabilità qualificata" dei minori nei confronti dell'intelligenza artificiale.

Partendo da questa premessa, abbiamo innanzitutto analizzato i rischi concreti ai quali i minori possono realisticamente essere esposti in tali ambienti. Dal punto di vista giuridico, è emerso chiaramente come le minacce riguardino, tra gli altri, il diritto alla protezione dei dati personali, alla dignità, al libero sviluppo della personalità, nonché i diritti alla privacy e all'immagine. I giocattoli intelligenti, infatti, raccolgono informazioni sugli utenti e sono in grado di memorizzare, elaborare e condividere dati relativi all'immagine, alla voce, all'attività fisica o alla geolocalizzazione tramite GPS, riferiti – vale la pena ribadirlo – a bambini in tenera età. In secondo luogo, gli algoritmi sono onnipresenti nei social network, in particolare attraverso i sistemi di raccomandazione e i feed, proponendo contenuti potenzialmente non adeguati alla fase evolutiva dell'utente; in questo contesto, i fornitori di servizi non si preoccupano né di verificare che gli utenti soddisfino i requisiti minimi di età, né di configurare gli algoritmi in modo tale da ridurre i rischi, preferendo invece soluzioni che ottimizzino i propri benefici. Infine, i deepfake stanno raggiungendo un livello di perfezionamento sempre più avanzato, risultando in grado di simulare chiunque nell'atto di compiere qualsiasi azione, riproducendo in modo sempre più fedele le caratteristiche fisiche e persino la voce della persona.

In conclusione, il presente contributo si propone come punto di partenza per avvicinare il lettore ai contesti realistici e concreti nei quali i minori risultano esposti ai rischi derivanti dall'intelligenza artificiale, in una prospettiva complementare rispetto a quegli studi di natura più teorica che analizzano i pericoli in termini astratti. A partire da tale base, sarà necessario interrogarsi sull'effettiva efficacia del GDPR, del DSA e del RIA nel garantire una tutela adeguata agli utenti, in particolare a coloro che, come i minori, appartengono a una categoria particolarmente vulnerabile, o se, al contrario, si renda opportuna una riflessione ulteriore su nuove forme di enforcement da rivolgere ai fornitori di servizi. Le valutazioni che, in questa direzione, saranno condotte una volta trascorso un ragionevole lasso di tempo dall'entrata in vigore dell'intero impianto normativo – il cui elemento più recente è rappresentato dal RIA – non sono, ad oggi, ancora disponibili.

Abstract

Negli ultimi due decenni abbiamo assistito all'inarrestabile ascesa di nuove tecnologie, in particolare quelle associate al Web 2.0 e ai social network. Oggi stiamo entrando in una nuova fase tecnologica caratterizzata dagli sviluppi dell'intelligenza artificiale (IA), inclusi big data, sistemi algoritmici e apprendimento automatico. Sebbene le opinioni accademiche varino dallo scetticismo all'estremo ottimismo sul potenziale dell'IA, non vi è dubbio che numerose applicazioni basate sull'IA siano già ampiamente utilizzate. Tra queste, giocattoli intelligenti, decisioni automatizzate, sistemi di raccomandazione e strumenti interattivi come ChatGPT. Come per tutti i progressi tecnologici, queste innovazioni comportano rischi per gli utenti, rischi particolarmente gravi per i minori. A causa della loro immaturità e del loro stadio di sviluppo, bambini e adolescenti potrebbero non possedere la consapevolezza necessaria per identificare e reagire ai pericoli di tali tecnologie. Questo articolo si concentra sui modi concreti in cui l'IA può rappresentare una minaccia reale per i minori, integrando analisi teoriche più ampie

con considerazioni pratiche e concrete.

Parole chiave: minori, intelligenza artificiale, regolamentazione digitale, smart toys, algoritmi, deepfakes

*

In the last two decades, we have witnessed the unstoppable rise of new technologies, particularly those associated with Web 2.0 and social networks. Today, we are entering a new technological phase marked by developments in artificial intelligence (AI), including big data, algorithmic systems, and machine learning. While scholarly views range from scepticism to extreme optimism about the potential of AI, there is no doubt that a number of AI-based applications are already widely used. These include smart toys, automated decisions, recommendation systems, and interactive tools like ChatGPT. As with all technological advances, these innovations carry risks for users—risks that are especially serious for minors. Due to their immaturity and developmental stage, children and adolescents may lack the necessary awareness to identify and respond to the dangers of such technologies. This article focuses on the concrete ways in which AI can pose real threats to minors, complementing broader theoretical analyses with practical, real-world considerations.

Keywords: minors, artificial intelligence, digital regulation, smart toys, algorithms, deepfakes