

**Il fenomeno del *Deepfake* e le sue derive**

*Avv.ta Graziella Dorotino*

Negli ultimi decenni l’Intelligenza Artificiale ha acquisito uno spazio sempre più ampio nella quotidianità di ciascuno di noi, basti pensare a determinati prodotti dell’industria cinematografica ( la trilogia di Avatar del regista James Cameron ne è un mero esempio), oppure, ad alcune applicazioni installabili sui dispositivi elettronici più comuni (smartphone, computer etc.) o ancora il significativo contributo che l’Intelligenza Artificiale sta fornendo alla scienza, in particolare, nel campo della medicina.

Si stima, infatti, che ad oggi l’80% dei dati utilizzati in ambito medico sia non strutturato, ossia generato per mezzo di algoritmi dell’Intelligenza Artificiale.

Dopo aver “istruito” una macchina, si immagini quelle adoperate per le tac, le risonanze magnetiche o per i semplici raggi X, è possibile identificare con un buon grado di affidabilità patologie tumorali, cardiovascolari, dermatologiche e respiratorie.[[1]](#footnote-1)

Sennonché, appurate tali migliorie, alcuni risvolti dell’Intelligenza Artificiale stanno destando non poca preoccupazione fra i giuristi, soprattutto alla luce delle risapute difficoltà per il legislatore di adattare, in breve tempo, la normativa alle nuove esigenze socioculturali.

Non di rado, infatti, le leggi appaiono lacunose ed insufficienti.

In particolare, nella presente dissertazione affronteremo l’assenza, nell’ordinamento italiano, di mezzi idonei ed utili a neutralizzare gli effetti del fenomeno cd. *Deepfake*.

Con tale termine si fa riferimento a video, foto ed audio creati utilizzando *software* di Intelligenza Artificiale i quali, partendo da contenuti reali di vario tipo (immagini o audio per l’appunto),modificandone o ricreandone in modo assolutamente realistico le caratteristiche, ne generano un corrispondente falso digitale.

L’espressione *Deepfake*, inoltre,rappresenta un neologismo dato dalla fusione dei termini *fake* (falso) e *deep learning* (una tecnologia dell’AI).[[2]](#footnote-2)

Ebbene, nei mesi il fenomeno poc’anzi descritto sta assumendo connotati che non consentono a chi, a vario titolo (forze dell’ordine, giudici, avvocati etc.), è chiamato ad intervenire ed inquadrare la corretta normativa applicabile.

Ad oggi le derive più preoccupanti sono essenzialmente due: il cd. *Phishing* *vocale* e il cd. *Deep nude.*

Con il primo termine, invero, si indica quell’operazione che consente di estrapolare la voce di una persona da video, audio o conversazioni dal vivo e, dopo di ciò, creare un nuovo *file* digitale con il suono acquisito.

A tal proposito, si è registrato un incremento non solo di truffe realizzate per mezzo di *file* prima generati nel modo sopra indicato e poi trasmessi con canali ordinari come WhatsApp e Telegram, ma anche episodi di cyberbullismo a danno principalmente di ragazzi adolescenti.

A seguito di numerose dichiarazioni raccolte, inoltre, è emerso come il *Phishing vocale* sia uno dei nuovi strumenti di ricatto adottati da compagni, mariti (o ex) abusanti nel ciclo della violenza domestica innescata.

Con il termine *Deep nude*, invece, si richiama quell’operazione informatica in conseguenza della quale persone possono essere rappresentate nude, in pose discinte, in situazioni compromettenti (si pensi in compagnia di un presunto amante) o perfino in contesti pornografici coinvolgenti anche minorenni.[[3]](#footnote-3)

Utilizzando appositi software di Intelligenza Artificiale, difatti, i visi di chiunque possono comparire su corpi di altri soggetti, nudi o impegnati in atti di natura esplicitamente sessuale, nonché si è reso possibile “spogliare” i corpi vestiti acquisiti da altre immagini, potendo creare immagini altamente realistiche.

In principio, episodi di questo tipo avvenivano prettamente a danno di personaggi famosi, quali attrici e cantanti (si ricordi la denuncia presentata dalla cantante statunitense Taylor Swift nel 2021 o quella presentata dalla rapper italiana Rose Villain, risalente a poche settimane addietro), probabilmente per la brama di ottenere dei soldi in cambio, oppure, sperando di ottenere attenzioni da parte della collettività attraverso la propria “creatività”.

Negli ultimi due anni, tuttavia, con la maggiore diffusione dei software che supportano questa tecnologia, il rischio di vedersi coinvolti in simili rappresentazioni interessa anche la gente comune.

Le prime denunce per diffusione di immagini di nudo create con l’Intelligenza Artificiale si sono verificate negli Stati Uniti dove, più di una decina di studentesse della Westfield High School del New Jersey si sono viste, loro malgrado, “spogliate” dai loro compagni sui gruppi WhatsApp della scuola.

Episodi di questo tenore hanno via via interessato sempre più distretti scolastici statunitensi come la Beverly Vista Middle School di Beverly Hills (California), il cui sovraintendente ha dovuto prendere contatti con la polizia locale dopo aver scoperto le immagini esplicite create con l’Intelligenza Artificiale e diffuse da cinque studenti della medesima scuola[[4]](#footnote-4).

Anche in Italia denunce di egual tipo non hanno tardato ad arrivare, a mero titolo esemplificativo, nell’aprile 2023, in una scuola media in provincia di Roma, due ragazzi sono stati accusati di aver modificato cinque foto delle compagne di classe con l’applicazione BikiniOff, al fine poi di renderle pubbliche.

Sono nell’aprile dello scorso anno sono state segnalate 27 infrazioni di questo tipo[[5]](#footnote-5).

Ancora.

Risale al luglio 2020 la diffusione in rete di alcune immagini *hard* ritraenti l’odierna Presidente del Consiglio, Giorgia Meloni, ad opera di due signori, padre e figlio residenti in Sardegna, i quali erano intenzionati a screditarne, in siffatto modo, la credibilità politica e personale.[[6]](#footnote-6)

Quanto sin qui riportato, è sufficiente per far riporre l’attenzione su di un particolare di non poco conto: fra le vittime del cd. Phishing vocale non risulta esservi prevalenza di uomini rispetto alle donne (o viceversa), di contro, la pratica di alterare e diffondere immagini rappresentative di nudi, annovera fra le sue principali vittime persone di sesso femminile.

Questa constatazione non può che intensificare la forte preoccupazione e lo sconforto provati poiché, nonostante le innumerevoli battaglie intraprese, c’è chi ancora (e non la minoranza purtroppo) strumentalizza la sfera sessuale delle donne per metterne in dubbio le capacità genitoriali, lavorative ed affettive; come se la rappresentazione di una donna nuda dovesse far ricordare di quanto il valore di una donna sia sempre relativo e direttamente collegato alle abitudini – anche sessuali – della medesima.

Solo da poche settimane, peraltro, è stata resa nota la notizia che a luglio 2024 si terrà la prima udienza del processo nei confronti dei due uomini che diffusero le immagini della Presidente del Consiglio e, altresì, che quest’ultima si sia costituita parte civile manifestando la volontà, in caso di vittoria, di voler devolvere quanto ottenuto a titolo di risarcimento del danno ad un fondo ministeriale che sostiene le donne vittime di violenza.

Questa scelta è apprezzabile per il sotto messaggio che, a parere di chi scrive, potrebbe esserne desunto: la minaccia di diffusione e la diffusione stessa di immagini sessualmente esplicite rappresentano una forma di violenza psicologica a tutti gli effetti, alimentata da una cultura misogina che vuole le donne etichettate in malo modo se coinvolte in contesti di tipo sessuale.

A ragion veduta, dunque, il Garante Nazionale per la Protezione dei dati personali esorta il legislatore ad intervenire sul testo dell’art. 612 ter c.p., disciplinante il delitto di *revenge porn* (introdotto con la L. 69/2019 nota come Codice Rosso), al fine di ampliare le ipotesi criminose riconducibili al sopracitato delitto[[7]](#footnote-7).

In fondo, occorrerebbe specificare come le immagini e video oggetto della diffusione illecita non debbano essere necessariamente reali ma anche frutto di operazioni informatiche.

Tale omissione, nel tempo, potrebbe causare una significativa lacuna normativa, soprattutto per chi pretende giustizia in sede penale, non essendo di facile individuazione la fattispecie criminosa di riferimento.

La condotta posta in essere dagli autori di *Deep nude*, invero, sembrerebbe presentare molte analogie con quanto punito ai sensi dell’art. 600 quater c.p., con il quale si assicura la configurabilità dei delitti di cui agli artt. 600 ter e 600 quater c.p. anche quando il materiale pornografico rappresenta immagini di minori degli anni 18 o parte di esse, di tipo virtuale.

Al comma II è chiarito, inoltre, come per immagini virtuali si intendano immagini realizzate con tecniche di elaborazione grafica non associate in tutto o in parte a situazioni reali.

Malgrado tali oggettive affinità strutturali, siamo ben consapevoli di come queste non siano sufficienti in ossequio al principio cardine del divieto di analogia in materia penale ex art. 14 delle preleggi e che, dunque, si faccia sempre più urgente la necessità di domandare al legislatore di intervenire per ostacolare un fenomeno sociale che, purtroppo, sta divenendo sempre più impattante nella sfera non solo privata, ma anche pubblica, delle persone.

Non a caso, difatti, molti sociologi hanno definito questo uso erroneo dei *software* di IntelligenzaArtificiale una “manovra di ingegneria sociale”, in ragione del fatto che la generazione di voci sintetiche ed anche la creazioni (e la successiva divulgazione) di immagini equivoche possono compromettere l’opinione pubblica su di una data personalità politica andando, di fatto, ad incidere sulla credibilità del suo programma elettorale.

Se tutto questo fosse vero, vorrebbe dire che l’uso distorto dell’Intelligenza Artificiale potrebbe causare – oltre al danno psicologico di cui ai paragrafi precedenti – una, seppur celata, violazione della libertà di pensiero, nell’accezione di libertà di ciascun individuo di crearsi un proprio convincimento su di un determinato aspetto politico, religioso, sociale etc.

Quanto detto si tradurrebbe, quindi, in un costante attacco alla Costituzione italiana (dove la libertà di espressione trova tutela all’art. 11) e, altresì, alle principali fonti di diritto sovranazionale che mirano a tutelare i diritti primari e le libertà di ciascun individuo (art. 10 Cedu, art. 11 Carta dei Diritti Fondamentali U.E. e art. 19 Dichiarazione Universale dei diritti umani).

1. Istituto di ricerche farmacologiche Mario Negri – IRCCS <https://www.marionegri.it/magazine/intelligenza-artificiale-medicina> . [↑](#footnote-ref-1)
2. Vademecum del Garante Nazionale per la Protezione dei dati personali anno 2022 <https://www.garanteprivacy.it/temi/intelligenza-artificiale/deepfake> . [↑](#footnote-ref-2)
3. Vademecum del Garante Nazionale per la Protezione dei dati personali anno 2023 <https://www.garanteprivacy.it/temi/intelligenza-artificiale/deepfake> . [↑](#footnote-ref-3)
4. New York Times <https://www.nytimes.com/2024/04/08/technology/deepfake-ai-nudes-westfield-high-school.html>. [↑](#footnote-ref-4)
5. Relazione EUROPOL 2023 <https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes> . [↑](#footnote-ref-5)
6. Wired <https://www.wired.it/article/giorgia-meloni-deepfake-porno-processo/> . [↑](#footnote-ref-6)
7. Vademecum del Garante Nazionale per la Protezione dei dati personali anno 2023 <https://www.garanteprivacy.it/temi/intelligenza-artificiale/deepfake>. [↑](#footnote-ref-7)