

FENOMENUL DEEFAKE

Av. Raluca ANDERCO

Abstract

Deepfake is a form of Artificial Intelligence (AI), whose algorithms are used to change people's faces from digital content and thus create a realistic-looking fake.

Deepfake content was originally only with celebrities, but even ordinary people can create their own deepfake content today. With the wide-spread use of deepfake content, issues such as audience manipulation, attacks on personal rights, violations of intellectual property rights and protection of personal data are becoming more common. Legislators and Big Tech are looking for an effective solution to the growing problem of deepfake.

In the field of advertising and media, this phenomenon has grown without limits, and the lack of a legal framework only expands the phenomenon even more.

From a legal perspective, the World Intellectual Property Organization ("WIPO") published the "Draft Issues Paper on Intellectual Property Policy and Artificial Intelligence" in December 2019.

The project contains deepfake content issues regarding intellectual property rights. There are two questions raised in connection with the deepfake issue in the project:

(i) Because deepfakes are created based on copyrightable data, who should copyright a deepfake?

(ii) Should there be a system of fair remuneration for people whose likenesses and "performances" are used in a deepfake?

WIPO concluded that deepfake can cause much more serious problems, such as the violation of human rights: the right to privacy, the right to self-image, the right to dignity, the right to privacy, the right to the protection of personal data, etc., by no means copyright infringement.

In the rest of this article, we will look at the legal options available to the victim of deepfake content.

Key words: *Deepfake, Artificial Intelligence, liability, human rights, GDPR, Civil code*

Deepfake – aspecte de ordin teoretic și practic

Deepfake este o formă a Inteligenței Artificiale (AI), ai cărei algoritmi¹ sunt folosiți pentru a schimba fețele oamenilor din conținuturile digitale și a crea, astfel,

¹ Algoritmii funcționează pe seturi masive de date – un deepfake destul de convingător se poate face cu 300 de imagini ce aparțin persoanei sursă. Apoi, algoritmi schimbă o față – cu o alta – cu tehnica de schimbare a feței. Cu toate acestea, există trei tehnici esențiale pentru a crea conținuturi Deepfake care se confruntă cu schimbul de expresii, rețelele adverse generative („GAN”). Indiferent de tehnica folosită, procesul are, în general, aceiași pași care sunt extracția, antrenamentul și crearea. De asemenea, nu mai este nevoie de seturi masive de date. Astăzi, chiar și o singură fotografie a unei surse este suficientă pentru a crea conținut deepfake.

un fals cu aspect realist. Dacă o persoană vrea să pună anumite cuvinte în gura unui personaj, acest lucru devine posibil prin crearea unui deepfake.

Conținutul deepfake era inițial doar cu celebrități, însă chiar și oamenii obișnuiți își pot crea propriul conținut deepfake astăzi. Odată cu utilizarea pe scară largă a conținutului deepfake, probleme precum manipularea publicului, atacurile la drepturile personale, încălcările drepturilor de proprietate intelectuală și protecția datelor cu caracter personal devin din ce în ce mai frecvente. Legislatorii și Big Tech² caută o soluție eficientă împotriva problemele tot mai mari ale deepfake-urilor.

În domeniul publicitar și media, acest fenomen a luat amploare fără margini, iar lipsa unui cadru legal, nu face decât să extindă și mai mult fenomenul³.

În 2017, un dezvoltator de software poreclit „deepfakes” pe platforma online Reddit și-a postat creațiile prin care a schimbat fețele celebrităților de la Hollywood pe fețele artiștilor porno. După ce creațiile s-au răspândit rapid, deepfake-ul a devenit o nouă tendință.

În consecință, au apărut numeroase aplicații pentru a crea conținuturi deepfake în ultimii 4 ani. Deoarece aplicațiile sunt ușor accesibile și gratuite, mulți oameni creează conținut deepfake despre ei înșiși sau despre celebrități, cum ar fi politicieni, actori/actrițe, artiști și cântăreți. Câteva exemple celebre de deepfake-uri sunt despre încoronarea Regelui Charles, Obama, Trump, Zuckerberg și Dalí, toate acestea destul de convingătoare. Cu alte cuvinte, este dificil să distingem care dintre ele este reală. Prin urmare, este emoționant să-l auzi pe artistul tău preferat strigându-te într-un videoclip sau să-i vezi pe actorii tăi preferați într-un alt film pe care nu l-au jucat.

Deși există numeroase probleme legate de deepfake, acest articol se va concentra pe aspectele juridice generate de fenomenul deepfake, având în vedere faptul că aceste drepturi civile nepatrimoniale sunt printre primele lezate de acest fenomen.

Deepfake – aspecte juridice

Organizația Mondială a Proprietății Intelectuale („OMPI”) a publicat „Proiectul de document cu probleme privind politica de proprietate intelectuală și inteligența artificială” în decembrie 2019.

² *Big Tech*, denumire cunoscută și sub numele de Tech Giants, Big Five sau S&P 5, reprezintă cele mai mari și dominante companii din industria tehnologiei informației din Statele Unite ale Americii, și anume Amazon, Apple, Facebook, Google și Microsoft.

³ Deepfake-urile au o utilizare largă, de la industria divertismentului până la sectorul sănătății. De exemplu, a fost folosit în sectorul sănătății pentru a detecta tumorile. Cu toate acestea, deepfake-urile au unele dezavantaje, cum ar fi politicile post-adevăr și hărțuirea femeilor sub formă de porno răzbunare (vom reveni asupra acestui aspect), manipularea publicului, intervenția la alegeri, încălcarea drepturilor de protecție a datelor cu caracter personal și a drepturilor de proprietate intelectuală.

Proiectul conține probleme de conținut deepfake în ceea ce privește drepturile de proprietate intelectuală. Există două întrebări ridicate în legătura cu problema deepfake în proiect:

„(i) Deoarece deepfake-urile sunt create pe baza datelor care pot face obiectul dreptului de autor, cui ar trebui să aparțină drepturile de autor asupra unui deepfake?

(ii) Ar trebui să existe un sistem de remunerare echitabilă pentru persoanele ale căror asemănări și „performanțe” sunt folosite într-un deepfake?

OMPI afirmă că deepfake-urile pot cauza probleme mult mai grave, cum ar fi încălcarea drepturilor omului: dreptul la confidențialitate, dreptul la propria imagine, dreptul la demnitate, dreptul la viață privată, dreptul la protecția datelor cu caracter personal⁴ etc., nicidecum încălcarea drepturilor de autor.

Prin urmare, potrivit OMPI, principala preocupare aici este dacă dreptul de autor ar trebui să fie acordat chiar și imaginilor deepfake?

OMPI afirmă că, dacă conținutul deepfake este complet contradictoriu cu viața victimei, conținutul acestuia nu ar trebui să fie recompensat cu protecția drepturilor de autor, ci ar trebui sancționat conform normelor juridice de drept civil, opinie la care ne raliem.

Pentru întrebările anterioare, OMPI mai menționează că, dacă deepfake-urile sunt supuse dreptului de autor, ar trebui să aparțină inventatorului deepfake-urilor, însă nu asta este, de fapt, problema juridică principală⁵.

Dreptul de autor nu este o armă juridică ce poate fi pusă împotriva deepfake-ului, deoarece victima deepfake-urilor nu deține un drept de autor în propria imagine.

Prin urmare, victima deepfake-ului poate folosi calea dreptului la protecția datelor cu caracter personal. Conform articolului 5 alineatul (1) din Regulamentul general al UE privind protecția datelor („GDPR”), „datele cu caracter personal trebuie să fie exacte și, acolo unde este necesar, păstrate la zi; trebuie luate toate măsurile rezonabile pentru a se asigura că datele cu caracter personal care sunt

⁴ Drepturi fundamentale, reglementate de orice constituție democratică.

⁵ Trebuie menționat că în România există deja un prim proces pe tema deepfake: O instanță din Giurgiu a judecat primul proces din țară cu tema „Deepfake”. Acuzata a fost o elevă care a făcut un trucaj video cu un coleg. Minora a fost obligată de instanță să achite 5.000 de lei daune morale. „Soluția pe scurt: Admite în parte acțiunea formulată... Obligă pârâții la plata sumei de 5000 de lei, cu titlu de daune morale (răspundere civilă delictuală a părintelui pentru fapta ilicită a minorului). Respinge în rest acțiunea. Admite în parte cererea reconvențională formulată de pârâții-reclamanți. Obligă reclamanții-pârâți la plata sumei de 5.000 de lei către pârâții-reclamanți în calitate de reprezentanți legali ai minorei, cu titlu de daune morale (răspundere civilă delictuală a părintelui pentru fapta ilicită a minorului). Respinge în rest cererea. Dispune compensarea cheltuielilor de judecată. Cu drept de apel, în termen de 30 de zile de la data comunicării. Cererea se va depune la Judecătoria Giurgiu”, se arată pe site-ul Judecătoriei Giurgiu.

inexacte, având în vedere scopurile pentru care sunt prelucrate, sunt șterse sau rectificate fără întârziere („exactitate”).

În lumina articolului 5, dacă un conținut deepfake este irelevant, inexact sau fals, acesta ar trebui șters sau rectificat fără întârziere.

Mai mult, chiar și dacă un conținutul deepfake este adevărat sau exact, o persoană vizată – victima unui deepfake – își poate exercita dreptul de a fi uitat, acordat rezidenților europeni în articolul 17 din GDPR ca „dreptul la ștergere”.

În conformitate cu articolul 17, persoana vizată are dreptul de a obține de la operator ștergerea datelor cu caracter personal care o privesc fără întârzieri nejustificate, iar operatorul de date are obligația de a șterge datele cu caracter personal fără întârziere.

Nu în ultimul rând, persoana fizică, victimă a Deepfake, are evident posibilitatea de a uzita și calea dreptului comun, care sancționează încălcarea drepturilor fundamentale ale persoanei fizice: dreptul la viață privată, dreptul la demnitate, dreptul la propria imagine, dreptul la confidențialitate.

Conform normelor juridice de drept, drepturile subiective nepatrimoniale enumerate anterior, sunt garantate și ocrotite.

Conform legislației din România, subiectul pasiv al unui deepfake are la dispoziție calea dreptului comun, reglementată de Codul civil la art. 71 – Dreptul la viață privată, art. 72 – Dreptul la demnitate, art. 73 – Dreptul la propria imagine, art. 74 – Atingeri aduse vieții private, art. 252-257 – Apărarea drepturilor nepatrimoniale.

Conform dispozițiilor art. 252 Cod Civil, orice persoană fizică are dreptul la ocrotirea valorilor intrinseci ființei umane, cum sunt viața, sănătatea, integritatea fizică și psihică, demnitatea, intimitatea vieții private, libertatea de conștiință, creația științifică, artistică, literară sau tehnică⁶.

Persoana fizică ale cărei drepturi nepatrimoniale au fost încălcate ori amenințate, poate cere oricând instanței:

- a) Interzicerea săvârșirii faptei ilicite, dacă aceasta este iminentă,
- b) Interzicerea încălcării și interzicerea pentru viitor, dacă aceasta durează încă;

⁶ Curtea Europeană a Drepturilor Omului a trasat conduite obligatorii care revin statelor în vederea protecției dreptului la viață și a surprins valențele cele mai complexe pe care le presupune protecția acestui drept, atât sub aspectul subiectiv, cât și din perspectiva procedurală.

Percepțiile noi din viața socială au dus la includerea în viața privată a unor aspecte moderne precum: avortul, homosexualitatea, tras-sexualitatea (CEDO, Christine Goodwin c. Regatului Unit, 11 iulie 2002), precum și anumite dezvoltări ale mijloacelor de comunicare, cum sunt interceptările telefonice sau ale corespondenței electronice, utilizarea bazelor de date personale informatizate.

Lipsa unei reglementări legale precise, care să determine cu exactitate sfera acelor date necesare identificării persoanelor fizice sau juridice utilizatoare, deschide posibilitatea unor abuzuri în activitatea de reținere. Preluare și utilizare date destinate publicului.

Cei care creează, interpretează, difuzează sau expun o operă de artă contribuie la schimbul de idei și opinii absolut necesar uneo societăți democratice.

c) Constatarea caracterului ilicit al faptei săvârșite, dacă tulburarea pe care a produs-o subzistă⁷.

Cel care suferă o încălcare a unor asemenea drepturi poate cere instanței să îl oblige pe autorul faptei să îndeplinească orice măsuri socotite necesare de către instanță pentru a ajunge la restabilirea dreptului său atins:

a) Obligarea autorului, pe cheltuiela sa, la publicarea hotărârii de condamnare;

b) Orice alte măsuri necesare pentru încetarea faptei ilicite sau pentru repararea prejudiciului cauzat.

De asemenea, persoana prejudiciată poate cere despăgubiri sau, după caz, o reparație patrimonială pentru prejudiciul, chiar nepatrimonial, ce i-a fost cauzat, dacă vătămarea este imputabilă autorului faptei prejudiciabile⁸.

Repararea prejudiciului este un principiu care rezultă din însăși esența răspunderii civile, care duce la înlăturarea tuturor consecințelor unui fapt dăunător și la restabilirea situației anterioare a celui păgubit sau lezat.

Repararea este necesar să se facă în natură, prin utilizarea mijloacelor celor mai potrivite realizării acestui scop. O asemenea finalitate nu poate fi obținută întotdeauna prin repararea bănească a prejudiciului, adică prin echivalent, ci numai prin repararea prin mijloace naturale, adecvate, firești.

Încetarea încălcării și interzicerea pentru viitor, dacă aceasta durează încă, presupune o faptă ilicită în desfășurare, atingerea dreptului producându-se deja (în formă continuă sau continuată). Spre exemplu, dacă atingerea dreptului la viață intimă constă în înregistrarea de imagini filmate fără acordul persoanei vătămate, instanța poate dispune distrugerea suportului magnetic pe care se află înregistrările, în cazul încălcării dreptului la imagine prin folosirea unor fotografii în pliante, calendare etc, se poate dispune retragerea de la distribuitori a materialelor menționate, în cazul folosirii fără drept a imaginii unor persoane pentru reclama la unele produse, se poate dispune scoaterea fragmentului incriminat din conținutul spotului publicitar etc.

Constatarea caracterului ilicit al faptei se va dispune după epuizarea acțiunii autorului atingerii ilicite, numai cu condiția ca tulburare pe care a produs-o fapta acestuia să subziste.

Deși succesiunea textelor celor trei litere ale alin. (1) stabilește o ordine de promovare a celor trei acțiuni, în funcție de momentul executării faptei ilicite și se poate concepe exercitarea separată, succesivă a acțiunilor, evident, nimic nu oprește transformarea unei acțiuni în alta, dacă desfășurarea activității ilicite este mai alertă decât procedura judiciară.

⁷ Art. 253 Cod civil.

⁸ Art. 253 alin. (4) Cod civil reprezintă practic petitul cererii de chemare în judecată pe care o poate formula victima unui conținut deepfake.

Aceste acțiuni în justiție sunt proceduri contencioase, desfășurate în ședință publică, cu respectarea principiului contradictorialității, cu administrarea întregului probatoriu. Deși drepturile personalității sunt opozabile *erga omnes*, în momentul săvârșirii faptei ilicite, subiectul ei pasiv se individualizează, devine determinat sau determinabil, astfel că va fi chemat în judecată autorul faptei ilicite.

Cât privește dreptul la viață privată, conform art. 71 Cod civil, nimeni nu poate fi supus vreunor imixțiuni în viața intimă, personală sau de familie (...) fără consimțământul său.

Este interzisă utilizarea, în orice mod, a informațiilor din viața privată, fără acordul acesteia sau fără respectarea limitelor prevăzute de lege.

În exercitarea dreptului la propria imagine, persoana fizică poate să împiedice reproducerea, în orice mod, a înfățișării sale fizice, a vocii sale sau utilizarea unei asemenea reproduceri.

Viața privată nu poate fi definită cu precizie, conținutul ei diferă în funcție de societatea în care individul trăiește sau grupul social căruia îi aparține, însă ea include, în mod cert, dreptul persoanei la viață intimă, personală, dreptul acesteia la viață socială și dreptul la un mediu înconjurător sănătos⁹.

Spre exemplu, intimitatea relațiilor personale constituie cadrul clasic al protecției vieții private. De asemenea, respectul vieții de familie impune statului obligația de a acționa astfel încât să permită celor interesați să ducă o viață familială normală și de a dezvolta relații afective în cadrul acesteia.

Nu în ultimul rând, protecția corespondenței constituie un alt aspect al vieții private, ocrotit de lege¹⁰.

Conform art. 72 din Codul Civil, orice persoană are dreptul la respectarea demnității sale. Este interzisă orice atingere adusă onoarei și reputației unei persoane, fără consimțământul acesteia ori fără respectarea limitelor prevăzute la art. 75 Cod Civil.

Dreptul la reputație și onoare al persoanei nu este consacrat în mod expres de către art. 8 din Convenție¹¹, însă protecția rezultă întrinsec, mai ales că acest drept este reglementat în alte acte normative.

Viața privată și reputația unei persoane sunt două concepte diferite. Cineva poate să încalce dreptul la viață privată al unei persoane, fără a aduce atingere

⁹ C. Birsan.

¹⁰ Art. 8 CEDO reglementează protecția vieții private, art. 26 din Constituția României reglementează același aspect, art. 33 din Decizia CNA 220/2011 și art. 89 din Legea nr. 8/1996 privind protecția drepturilor de autor și a altor drepturi conexe, toate aceste acte normative reglementează și ocrotesc dreptul la viață privată.

¹¹ C.E.D.O., Gunnarsson c. Islandei, 20 octombrie 2005, www.jurisprudencedo.com.

reputației acesteia, după cum același comportament poate aduce atingere ambelor aspecte ale vieții sale¹².

Demnitatea este greu de definit, dar presupune, în orice caz, cinste, bună-credință și reputație neatinsă. Orice atingere adusă demnității omului îl expune pe acesta excluderii din sfera relațiilor sociale. Demnitatea este lezată și atunci când anumite *exhibiții* sunt transpuse prin scrieri sau imagini TV/media, cu ignorarea respectului datorat pudorii publicului.

Noțiunea de onoare are trei sensuri: a) vizează o persoană care se distinge prin funcția sau meritele sale, ne evocă un semn excepțional de considerație, de stimă, de omagiu adus valorii sale; b) poate viza element al patrimoniului moral al unei persoane, care îl îndreptățește să se facă respectat de ceilalți și de a-i respecta pe aceștia; c) într-un sens obiectiv, înseamnă regulă de conduită, normă morală de comportament, valoare ideală la care se raportează o colectivitate sau grup.

Nu în ultimul rând, în lumina subiectului abordat în acest articol, este extrem de importantă reglementarea din Codul civil (art. 73) a Dreptului la imagine – în exercitarea dreptului la propria imagine, persoana poate să interzică ori să împiedice reproducerea, în orice mod, a înfățișării sale fizice ori a vocii sale sau, după caz, utilizarea unei asemenea reproduceri.

Această componentă a vieții private are ca scop protejarea identității persoanei, a sferei sale intime, a relațiilor sale personale, a libertății sexuale. Dreptul la propria imagine, care la origine a fost creația jurisprudenței, este un drept al personalității, un drept primordial pentru fiecare persoană. Această opțiune a persoanei de a cenzura folosirea imaginii sale se fundamentează pe libertatea și intimitatea sa, precum și nevoia de securitate.

Reproducerea unei persoane printr-o fotografie, film, desen, rețea internet etc. necesită acordul acelei persoane, consimțământul acordat unei persoane nu înseamnă în mod obligatoriu că este acordat și altora. În toate cazurile, consimțământul se prezumă a fi dat pentru captarea și difuzarea imaginii corecte a persoanei, nicidecum o reprezentare deformată.

Publicarea fotografiei unei persoane, fără consimțământul acesteia constituie o ingerință în dreptul la viața privată, fiind cu atât mai semnificativă atunci când respectiva persoană nu este cunoscută de către public¹³.

Iată cât de armonios se împletesc în continuare dispozițiile art. 71, 72 și 73 din Codul civil, cu dispozițiile art. 74 Cod Civil care reglementează atingerile aduse vieții private, pe care legiuitorul le enumeră, după cum urmează:

a) intrarea sau rămânerea fără drept în locuință sau luarea din aceasta a oricărui obiect fără acordul celui care o ocupă în mod legal;

¹² Ventola c. Banque Nationale du Canada, C.S., 5.11.1997.

¹³ CEDO, Sciacca c. Italiei, 11.01.2005.

b) interceptarea fără drept a unei convorbiri private, săvârșită prin orice mijloace tehnice, sau utilizarea, în cunoștință de cauză, a unei asemenea interceptări;

c) captarea ori utilizarea imaginii sau a vocii unei persoane aflate într-un spațiu privat, fără acordul acesteia;

d) difuzarea de imagini care prezintă interioare ale unui spațiu privat, fără acordul celui care îl ocupă în mod legal;

e) ținerea vieții private sub observație, prin orice mijloace, în afară de cazurile prevăzute expres de lege;

f) difuzarea de știri, debateri, anchete sau de reportaje scrise ori audiovizuale privind viața intimă, personală sau de familie, fără acordul persoanei în cauză;

g) difuzarea de materiale conținând imagini privind o persoană aflată la tratament în unitățile de asistență medicală, precum și a datelor cu caracter personal privind starea de sănătate, problemele de diagnostic, prognostic, tratament, circumstanțe în legătură cu boala și cu alte diverse fapte, inclusiv rezultatul autopsiei, fără acordul persoanei în cauză, iar în cazul în care aceasta este decedată, fără acordul familiei sau al persoanelor îndreptățite;

h) utilizarea, cu rea-credință, a numelui, imaginii, vocii sau asemănării cu o altă persoană;

i) difuzarea sau utilizarea corespondenței, manuscriselor ori a altor documente personale, inclusiv a datelor privind domiciliul, reședința, precum și numerele de telefon ale unei persoane sau ale membrilor familiei sale, fără acordul persoanei căreia acestea îi aparțin sau care, după caz, are dreptul de a dispune de ele.

Se poate observa, din lecturarea textelor de lege anterior faptul că nu există niciun fel de clemență pentru folosirea, fără acord a imaginii unei persoane, victima fenomenului deepfake având la dispoziție nu doar temeiul juridic al acțiunii sale, dar și o procedură pentru sancționarea persoanei responsabile și de înlăturare a conținutului deepfake din media.

Concluzie

Astăzi, unele state și BigTech-urile iau măsuri împotriva problemei deepfake. De exemplu, BigTechs dezvoltă instrumente pentru a detecta provocările conținutului deepfake. Virginia, Texas și California sunt primele state din SUA care au reglementări împotriva problemelor legale ale deepfake-urilor. Legea din Virginia impune sancțiuni penale pentru distribuirea de pornografie deepfake neconsensuală, în timp ce legea din Texas interzice crearea și distribuirea de videoclipuri deepfake destinate să dăuneze candidaților la funcții publice sau să

influențeze alegerile. Când se iau în considerare problemele actuale, probabil că în curând vor apărea mai multe legislații și instrumente tehnologice împotriva deepfake-urilor.

România este departe în acest proces legislativ, chiar dacă există un cadrul legal general, acesta este totuși lipsit de o utilitate rapidă și imediată, iar industria media și casele mari de producție folosesc în continuare nestingherite acest procedeu.

Cred că ar fi necesar, în contextul dinamicii acestui fenomen, reglementarea unei proceduri urgente, cu efect rapid și cu sancțiuni mult mai aspre, pentru combaterea acestui fenomen.