

DEEPPFAKE - FABRICAREA REALITĂȚII CU AJUTORUL INTELIGENȚEI ARTIFICIALE

Laura TUGAREV

Doctor în științe ale comunicării, lector universitar, Universitatea de Stat din Moldova,
Chișinău, Republica Moldova,

e-mail: tugarevlaura@gmail.com

<https://orcid.org/0000-0002-1304-7209>

Dezvoltarea tehnologiilor informaționale au formatat parametrii realității. Pătrunse în diverse sfere ale activității umane, acestea schimbă modul de percepere, de interacțiune, de socializare dintre indivizi și grupurile sociale și chiar produc modificări în câmpul comportamental și cel acțional al omului. Prezentul articol reliefează valențele fenomenului deepfake, care a pătruns în spațiul informațional virtual și a câpătat extensiune în ultimii ani. Pe măsură ce tehnologia inteligenței artificiale este folosită în modificarea sau/și fabricarea realității, devine tot mai dificil de a identifica manipularea, dezinformarea în mediul online. În această ordine de idei, este imperios de a cerceta specificitatea acestui concept și totodată, de a cartografia un set de acțiuni care ar preveni și ar diminua acest fenomen.

Cuvinte-cheie: deepfake, new media, internet, inteligență artificială, conținut fals.

DEEPPFAKE - MANUFACTURING THE REALITY WITH THE AID OF THE ARTIFICIAL INTELLIGENCE

The development of information technologies has formed the parameters of reality. Penetrated into various spheres of human activity, they change the way we perceive, interact, socialize between individuals and social groups and even produce changes in the behavioral and acting field of man. This article highlights the meanings of the deepfake phenomenon, which has penetrated the virtual information space and gained extension in recent years. As artificial intelligence technology is used in modifying and / or fabricating reality, it becomes increasingly difficult to identify manipulation, misinformation in the online environment. In this regard, it is imperative to investigate the specificity of this concept and at the same time to map a set of actions that would prevent and diminish this phenomenon.

Keywords: deepfake, new media, internet, artificial intelligence, false content.

DEEPPFAKE - FABRICATION DE LA RÉALITÉ AVEC L'INTELLIGENCE ARTIFICIELLE

Le développement des technologies de l'information a formaté les paramètres de la réalité. Pénétrés dans diverses sphères de l'activité humaine, ils modifient le mode de perception, d'interaction, de socialisation entre les individus et les groupes sociaux et produisent même des changements dans le champ comportemental et actionnel de l'homme. Cet article met en évidence les valences du phénomène deepfake, qui a pénétré l'espace virtuel de l'information et s'est étendu ces dernières années. Comme la technologie de l'intelligence artificielle est utilisée pour modifier ou / et fabriquer la réalité, il devient de plus en plus difficile d'identifier la manipulation, la désinformation dans l'environnement en ligne. Dans cette optique, il est impératif de rechercher la spécificité de ce concept et en même temps de cartographier un ensemble d'actions qui permettraient de prévenir et de diminuer ce phénomène.

Mots-clés: deepfake, nouveaux médias, internet, intelligence artificielle, faux contenu.

DEEPFAKE – СОЗДАНИЕ РЕАЛЬНОСТИ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Развитие информационных технологий перестроило параметры реальности. Проникая в различные сферы человеческой деятельности, они меняют способ восприятия, взаимодействия, социализации между индивидами и социальными группами и даже производят изменения в поведенческом и деятельностном поле человека. В данной статье освещаются валентности явления дипфейк, вошедшего в виртуальное информационное пространство и получившего распространение в последние годы. Поскольку технология искусственного интеллекта используется для изменения и/или формирования реальности, становится все труднее выявлять манипуляции и дезинформацию в Интернете. В связи с этим, необходимо исследовать специфику данного понятия и одновременно наметить комплекс действий, которые предотвратили бы и уменьшили это явление.

Ключевые слова: дипфейк, новые медиа, Интернет, искусственный интеллект, фейковый контент.

Introducere

Dezvoltarea tehnologiilor informaționale au formatat parametrii realității. Pătrunse în diverse sfere ale activității umane, acestea schimbă modul de percepere, de interacțiune, de socializare dintre indivizi și grupurile sociale și chiar produc modificări în câmpul comportamental și cel de activitate al omului [1]. În prezent, progresul tehnologic este marcat de evoluția fără precedent a tot ceea ce înseamnă inteligența artificială (AI) - asistenți virtuali, programe de analiză a imaginilor, motoare de căutare, sisteme de recunoaștere vocală și facială. Inteligența artificială este considerată un element central al transformării digitale, astfel devenind o prioritate pentru societăți, sisteme, instituții.

Așadar, „inteligența artificială, ca urmare, creează noi posibilități pentru dezvoltarea umană. Deschide noi posibilități de dezvoltare, sporind eficiența activităților umane, inclusiv cea economică. ... concomitent cu impactul pozitiv, inteligența artificială produce și efecte considerabile negative. Astfel, atitudinea față de utilizarea inteligenței artificiale în diferite activități umane este duplicitară – este benefică utilizarea acesteia în scopuri umane, dar folosirea ei include și pericole pentru omenire” [2, p. 38]. Un așa „pericol” care a apărut în cadrul internațional mai întâi, este *deepfake*-ul – un produs sintetic, con-

ceput și realizat cu ajutorul sistemelor inteligente ce conține valențe de *deep learning* (învățare profundă) și conținut *fake* (fals). *Deepfake*-ul este o tehnologie bazată pe algoritmi IA, cu scopul de a modifica conținutul original, autentic, generând produse audio sau video asemănătoare, pe care individul le crede adevărate.

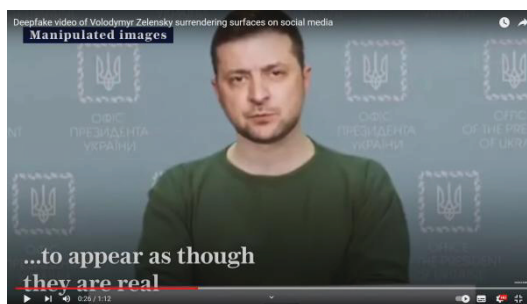
Context

„La nivel internațional, discuțiile despre *deepfake* au început din 2016, însă la sfârșitul anului 2017 acestea capătă amploare, după ce un utilizator cu *nickname*-ul *Deepfakes* a creat și a plasat în spațiul virtual video-uri în care fețele unor simpli utilizatori erau înlocuite cu cele ale unor actori celebri și/ sau ale unor actori transformați instant în colegii lor” [3].

Așadar, *deepfake*-ul este tehnologia folosită cu scopul de a produce sau de a modifica digital anumite informații: imagini, conținut audio sau video prin aplicarea algoritmilor IA, care au un înalt indicator de înșelăciune, dar sunt create în așa fel, încât acestea în percepția consumatorului sunt autentice și veritabile, „tehnologia are la bază un concept numit *Generative Adversarial Network* ce poate fi folosit pentru a reproduce imagini video și sunet superficial-autentice pentru ochiul uman” [4].

De obicei, *deepfake*-ul este folosit în sfera politică sau în cea a divertismentului. Politicienilor, actorilor, artiștilor etc., în timp real sau înregistrat, le pot fi modificate următoarele: discursul, mimica, limbajul non-verbal etc. „*Deepfake*-ul schimbă atributele (de exemplu: față, voce, tonul pielii, sexul, culorile, accesorii...) unei surse cu cele ale unei ținte” [5, pp. 1-17]. Așadar, mișcări artificiale ale corpului, modularea și vibrațiile vocii puțin schimbate, mișcarea nenaturală a ochilor sau nemișcarea (lipsa clipitului), expresii faciale nefirești, adăugarea unor elemente precum: păr, accesorii și altele, toate acestea conduc la existența unor premise în identificarea un *deepfake*.

Ținta *deepfake*-urilor au fost mai mulți liderii politici cum ar fi: Barack Obama, Donald Trump, Vladimir Putin, Angela Merkel, Vladimir Zelensky etc., dar și exponenți din lumea divertismentului: Tom Cruise, Will Smith etc.



Sursa: <https://www.youtube.com/watch?v=X17yrEV5sl4&t=10s>

Deepfake-ul este un produs al tehnologiilor, cu un impact major în modelarea și manevrarea unor curente ale opiniei publice. Cătălin Pătrașcu, expert în securitate cibernetică menționa despre *deepfake* că „e o alternativă complexă a *fake-news*-ului, o evoluție a *fake-news*-ului, un *next level*. E consecința tehnologiei” [6].

Evoluția tehnologiilor moderne și ale derivatelor sale au înlesnit apariția diverselor aplicații și instru-

mente care susțin indivizii în crearea rapidă a unui *deepfake*, cum ar fi de exemplu, aplicația *Zao* „care le permite utilizatorilor să-și adauge fețele la o listă de personaje de televiziune și filme pe care sistemul s-a antrenat. Un alt exemplu este platforma care permite utilizatorilor să își producă propriile video după principiile *deepfake*. Această companie își numește produsul tehnologie *deepfake* responsabilă, menționând că fiecare defect generat pe instrumentul lor are un filigran clar vizibil, care indică faptul că videoclipul este un *deepfake*” [7]. O altă aplicație *Deepfakes* este un software online care funcționează în cloud și creează videoclipuri *deepfake*.

FaceApp, *Faceswap*, *Face2Face*, *Facecrop* sau *Lyrebird* pot ajuta individul în intențiile sale de a schimba unele elemente dintr-o fotografie sau dintr-un produs video, cum ar fi: vârsta, culoarea tenului, culoarea sau lungimea părului, luminozitatea, saturația, mărimea ochilor, buzelor etc., totodată, aceste programele pentru Android sau IOS pun la dispoziție și diferite filtre care modifică varianta inițială a imaginii, nu vorbim neapărat despre un *deepfake*, ci despre posibilitățile virtuale care fac posibilă editarea, retușarea, trucarea sau fabricarea unor particularități, însușiri etc., care nu corespund cu realitatea.

Mediul online a devenit un „habitat” pentru *deepfake*-uri. Rețelele de socializare cum ar fi: Facebook, Twitter, Instagram, Tik Tok, Flickr, Snapchat or, platforme precum Youtube au devenit medii favorabile, garantând un circuit informațional sporit și tirajare online în timp real.

Generațiile emergente vor trebui să fie din ce în ce mai pregătite pentru detectarea *deepfake*-urilor, deoarece acestea sunt tot mai frecvente, procesul va fi unul similar cu cel al detectării știrilor false.

Sintetizând, putem enumera o serie de indicatori care reflectă transversal particularitățile unui *deepfake*:

a) este un conținut sintetic, creat în baza unor algoritmi ale inteligenței artificiale;

b) este un produs al tehnologiei, aflându-se la un nivel avansat față de *fake news*;

c) este o tehnică de sintetizare a imaginilor și sunetelor pentru a crea conținut care pare a fi adevărat, deși nu este;

d) este un instrument antrenat în procesul de dezinformare;

e) este mai greu de identificat în comparație cu știrile false, deoarece conținutul livrat este aproape identic cu cel real.

Pe 4 august 2021, în presa din Republica Moldova este publicat un material în exclusivitate, marca Jurnal TV, care urma să fie distribuit în spațiul informațional până la data alegerilor prezidențiale. În video Maia Sandu primește un plic de la Ilan Șor, protagonistă este creată cu ajutorul IA [8]. Pentru prima oară, o pleiadă de specialiști au menționat că este vorba despre un deepfake, iar alții au afirmat că este doar un conținut trucat, fals, care creează impresia unui material autentic.



Sursa: <https://www.youtube.com/watch?v=qzgLxAfPFTs>

În anul 2022, la începutul lunii octombrie, în mediul online apare o înregistrare a președintei Maia Sandu într-o pretinsă discuție cu prim-ministrul ucrainean, Denis Șmihal. Pe 3 octombrie 2022, Administrația Prezidențială dezmente autenticitatea declarațiilor atribuite Președintei Republicii Mol-

dova, iar în comunicatul de presă este menționat că „Înregistrarea video este o tentativă de implicare a Președintei Maia Sandu într-o manipulare de tip *deepfake*. Declarațiile prezentate în această falsificare nu corespund realității, fiind un trucaj prin montare de imagini și sunet” [9]. Așadar, *deepfake*-ul este un fenomen periculos, întrucât poate fi un instrument în strategiile de denigrarea și defaimare a imaginii unei persoane publice sau a unui stat. Cercetătorul Constantin Marin nota că „un clip deepfake poate avea un puternic impact vizual și emoțional, poate influența opinia publică și poate fi folosit pentru răzbunare sau defaimarea persoanelor publice” [10, p. 276].

Identificarea conținutului de *deepfake*, în fiecare zi devine tot mai greu de realizat, deoarece tehnologia progresează rapid, iar instituțiile abilitate nu reușesc să proiecteze metode, tehnici, programe, aplicații specifice. Pe de o parte, avem persoane afiliate, specializate pe această dimensiune, care creează deepfake-uri având un scop determinat, iar pe de altă parte, se reliefează încă o parte a tabloului, odată ce la îndemâna fiecărui individ se află diverse aplicații accesibile și gratuite, monitorizarea și detectarea devine practic imposibilă, iar fluxul generat zilnic în spațiul virtual devine incontrollabil.

Totuși, putem evidenția câteva direcții/acțiuni care ar spori gradul de combatere a conținutului *deepfake*:

1. Îmbunătățirea legislației pe plan internațional și implementarea unui cadru legal la nivel național care ar veni cu anumite sancțiuni în acest sens.

2. Dezvoltarea unor software-uri/aplicații care ar identifica *deepfake*-urile din spațiul informațional virtual.

3. Promovarea educației mediatice care ar contribui la dezvoltarea competențelor de analiză și interpretare critică a diferitor tipuri de produse informaționale.

În aceeași ordine de idei, este importantă și verificarea faptelor (*fact-checking*), „care constă în stabilirea autenticității unor informații prin confruntarea lor cu faptele, cu dovezile și cunoștințele disponibile” [11, p. 99]. La nivel internațional deja s-au coagulat eforturile în vederea creării unor site-uri care ar permite detectarea conținuturi generate de AI. Spre exemplu, site-ul *DetectFakes* – pune la dispoziția publicului mai multe a) aspecte conceptuale despre *deepfake*-uri și b) exemple care îl ajută pe consumatorul de informații să discearnă realul de fals. Un alt exemplu este *Digital Future Society* (program susținut de Ministerul Afacerilor Economice și Transformării Digitale al Guvernului Spaniei, în colaborare cu Mobile World Capital Barcelona) care a inițiat un proiect pilot, faza de execuție ianuarie 2022 - ianuarie 2023. *DFDLab* are ca scop prototiparea unor soluții bazate pe web care să sprijine utilizatorii în sesizarea și perceperea specificului, să aibă acces la o bază de date de *deepfake* și la două instrumente de verificare care au fost dezvoltate de Centrul pentru Cercetări și Tehnologie Hellas.

Concluzie

Epoca digitală încă impulsionează dezvoltarea tehnologiilor, individul se află într-o conexiune permanentă cu mediul online. Evoluția inteligenței artificiale va intensifica această legătură cu dispozitivele și mijloacele tehnice, iar computerul și internetul ar putea să devină surse primare din care oamenii se vor documenta, de aceea, este important ca „utilizarea noilor tehnologii să fie făcută cu atenție” [12, p. 199], ca instruirea pe filieră media și pe cea digitală să devină un imperativ, într-o epocă în care ne sunt difuzate pe lângă informații veridice, obiective și diverse pseudo-adevăruri.

Mulți experți în domeniul AI și securitate cibernetică susțin că *deepfake*-urile vor deveni mult mai

sofisticate proporțional cu dezvoltarea tehnologică în viitor. În acest sens, ar putea apărea amenințări mai grave pentru guverne și pentru opinia publică, deoarece acestea sunt capabile să creeze interferențe în competițiile electorale, tensiuni politice și activități criminale.

Referințe bibliografice

1. MORARU, Victor; PINTILII, Ionel. *Internetul, ca „motor al schimbărilor”, în vizorul cercetării științifice*. In: Akademos. Revista de Știință, Inovare, Cultură și Artă, 2022, nr. 3(66), pp. 121-127.
2. CIOBANU, Cesclav, CAPSIZU, Valeriu. *Inteligența artificială – o nouă etapă în dezvoltare*. Disponibil: https://ibn.idsi.md/sites/default/files/imag_file/37-47_3.pdf (vizitat 20.04.23)
3. TACU, Mariana. *Deepfake-ul vs inteligența artificială: ultima frontieră tehnologică în materie de știri false*. Disponibil: <https://old.mediacritica.md/ro/deepfake-ul-vs-inteligența-artificială-ultima-frontiera-tehnologică-materie-de-stiri-false/> (vizitat 20.04.23)
4. TRUȚĂ, Filip. *Ce e fenomenul deepfake și cum ar putea afecta România*. Disponibil: <https://romania.europalibera.org/a/ce-e-fenomenul-deepfake-si-cum-ar-putea-afecta-romania/30199564.html> (vizitat 24.04.23)
5. CAMPBELL Colin, PLANGGER, Kirk., SANDS, Sean., KIETZMANN, Jan. *Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising*. In: Journal of Advertising, 2022, Vol. 51, Issue 1, pp. 22-38.
6. NAGY, Julia, PĂTRAȘCU, Cătălin. *Fenomenul „deepfake” – cutia tech a Pandorei a fost deschisă*. Disponibil: <https://mindcraftstories.ro/tehnologie/fenomenul-deepfake-cutia-tech-a-pandorei-a-fost-deschisă/> (vizitat 30.04.23)
7. *Ce este deepfake?* Disponibil: <https://mediacritica.md/podcast/ce-este-deepfake/> (vizitat 30.04.23)
8. Jurnal TV - *Imagini fake cu Maia Sandu și Ilan Șor*. Disponibil: <https://www.youtube.com/watch?v=qzgLxAfPFTs> (vizitat 30.04.23)
9. Președinția dezmințe autenticitatea declarațiilor

atribuite șefei statului. Președinția Republicii Moldova [Site]. Publicat 3 octombrie 2022.. Disponibil: <https://presedinte.md/rom/presa/presedintia-dezminte-autenticitatea-declaratiilor-atribuite-sefei-statului> (vizitat 30.04.23)

10. MARIN, Constantin. *Mass-media în postcomu-*

nism. Cazul Republicii Moldova. Chișinău: CEP USM, 2021

11. CORAVU, Robert. *Evaluarea informațiilor din domeniul online*. București: Editura Universității din București, 2021.

12. MORARU, Victor (coord.). *Europenizarea: fațetele procesului*. Chișinău: IIESP, 2013.

