37. Országos Tudományos Diákköri Konferencia
Állam- és Jogtudományi Szekció

Budapest, 2025. március 24–26.

Rendező intézmény:
Károli Gáspár Református Egyetem
Állam- és Jogtudományi Kar

Deepfake: A modern büntetőeljárás legnagyobb kihívása

Kézirat lezárásának dátuma:

2024. december 10.

Tartalom

1. Bevezetés. A kutatás módszertana. 3
2. A deepfake fogalma és rövid történeti áttekintése 4
2.1. A deepfake fogalma 4
2.2. Történeti áttekintés 7
2.3. Deepfake alkalmazások és a mögöttük rejlő pszichológia 9
3. A mélyhamisítás kockázatai és az abból származó hátrányok 11
3.1. Kihatás a demokratikus rendszerekre 12
3.2. Az egyéneket érintő károk 13
3.3. A közvélemény befolyásolása és a politikai folyamatok torzítása 15
3.4. Nemzetközi kapcsolatok stabilitásának megingatása 18
4. A deepfake pozitív oldalai és a felmerülő szerzői jogok kérdése 19
4.1. A deepfake előnyei 19
4.2. Az általános adatvédelem követelménye 21
4.3. Az önkifejezés szabadsága 22
5. A deepfake jelenlegi jogi szabályozása, különös tekintettel az EU szabályokra 23
6. A deepfake hatása a büntető eljárásjogra 27
6.1. Kihívások a bíróságon: bizonyíték, védelem és bírói szkepticizmus 28
6.2. Javaslatok a büntetőeljárás során felmerülő problémák megoldására 29
7. A jogi szabályozás további lehetőségei és a felelősség megállapítása 31
8. Összegzés 36
9. Felhasznált irodalom 37
9.1. Szakirodalom 37
9.2. Internetes források 38
9.3 Felhasznált képek 39

1. Bevezetés. A kutatás módszertana.

Felgyorsult világunkban kép- és hanganyagok tömkelege áraszt el minket nap mint nap. Az információáradat valóságos iránytűként hat a hétköznapokban, amely segít megfigyelni és megfejteni a nagyvilágban végbemenő, másodpercről másodpercre változó folyamatokat, legyen szó politikáról, gazdaságról, időjárásról, szórakozásról. Egy pillanat alatt hozzáférhetünk a szükséges adatokhoz, amelyekkel továbbfűzhető a végtelen, életünket behálózó hírláncolat.

A technológia rohamos fejlődésével azonban olyan tartalmak jelennek meg, amik elferdítik a valóságot vagy akár nem is felelnek meg annak, mivel teljes mértékben hamisak. Az, hogy mi áll emögött, szinte mindenki számára nyilvánvaló. A mesterséges intelligencia (a továbbiakban: MI, angolul: artificial intelligence, AI) ugyanis egy olyan jelenség a 21. században, amiről csak kevesen nem hallottak. Hatásai, előnyei, hátrányai, illetve a benne rejlő potenciális veszélyek azonban közel sem olyan egyértelműek, mint maga a létezése, még a hozzáértőknek sem, nemhogy a laikusoknak. Márpedig használata mindennapossá vált, így ezzel párhuzamosan egyre csak növekszik az MI-vel való visszaélések és a vele összefüggésbe hozható bűncselekmények száma. Felmerülhet a kérdés, hogy miért fontos ezzel kapcsolatban olyan sok kutatást végezni, az ehhez kapcsolódó jogi intézkedéseket sürgetni, minél több ember figyelmét felhívni? Nos, az MI-vel összefüggésbe hozható tevékenységek (főként videók) száma az utóbbi időben 900%-kal nőtt[1], amely adat igencsak kétségbeejtő, főleg, ha tekintettel vagyunk arra a tényre, hogy az információtechnológia fejlődésével párhuzamosan az MI is fejlődik, amivel egyre könnyebbé válik a deepfake tartalmak gyártása. Ezt tetézve az online térben a készítők könnyedén leplezhetik magukat, főként, ha VPN-t, azaz virtuális magánhálózatot[2] használnak, amely által a legtöbb esetben lekövethetetlen, ki és milyen módon szerkesztette, terjesztette az adott deepfake terméket.

A deepfake-ek terjedésének korlátozására a legalapvetőbb megoldásnak talán a jogi szabályozás tűnhet, ám a jelenlegi jogi normák nem célirányosak az MI jelenségét és abból származó problémákat tekintve. Az idő azonban szorít, és fel kell vele venni a versenyt, ha szeretnénk megállítani a káros hatások végtelenségig tartó burjánzását.

[1] https://bigthink.com/the-present/reality-vs-deepfake/

[2] Neve onnan ered, hogy a VPN-en keresztül érkező és kimenő adatok a titkosítás miatt nem láthatók az eredeti hálózaton, így az adatforgalom részleteit csak a VPN-szolgáltató látja.

A dolgozat a deepfake technológia büntetőeljárásra gyakorolt hatásait elemzi. A kutatás alapját kvalitatív elemzések képezik, amelyek a meglévő szakirodalom, jogi szabályozások, illetve az álhírek és a manipulált audiovizuális tartalmak társadalomra gyakorolt hatásainak feltárásán alapulnak. A dolgozat a deepfake technológia fejlődését és annak kihívásait különféle szakterületeken vizsgálja, valamint kiemelten foglalkozik a technológia kriminalisztikai alkalmazhatóságával és a bizonyítékok hitelességének kérdésével a büntetőeljárásban. A dolgozatban alkalmazott módszerek között a nemzetközi esettanulmányok áttekintése is szerepel, különös tekintettel arra, hogy a deepfake technológia hogyan befolyásolja a jogi szabályozást, és milyen eszközökkel próbálkoznak a jogalkotók a fenyegetések kezelésére.

Hipotézisem az volt, hogy a deepfake széles körben való terjedése a büntetőeljárásban alapvető bizalmi problémákat idézhet elő, mivel a bírák (még akár a szakértők is) hajlamosak lehetnek a „seeing is believing" szemléletmódra, vagyis a vizuálisan megtapasztalt dolgok valóságként való elfogadására. A kutatás egyik alapvetése, hogy a deepfake megkérdőjelezi a digitális bizonyítékok hitelességét, ami hosszú távon a büntetőeljárásba vetett közbizalmat is aláásha tja, és még a valós, megbízható bizonyítékokkal szemben is szkeptikussá válhatunk.

Szerzőként nyilatkozom, hogy a dolgozat összeállítása során nem használtam mesterséges intelligencia-eszközt.

2. A deepfake fogalma és rövid történeti áttekintése

2.1. A deepfake fogalma

Annak érdekében, hogy mélyebben megértsük a témakört, először is tisztáznunk kell, mit jelent a deepfake. A deepfake olyan kép, hang vagy más anyag, amely részben vagy teljes egészében hamisított, vagy olyan létező kép, hang, egyéb anyag, amelyet kifinomult technikai eszközökkel manipuláltak, és amely aligha megkülönböztethető a valóditól (már ha megkülönböztethető)[3]. A deepfake-ek – vagy magyarra lefordítva „mélyhamisítások" – az ún. gépi tanulás (machine learning)[4] és az MI által jönnek létre. Amellett, hogy bizonyos mintákat gyártanak le adott metódusok segítségével, ezen hálózatok könnyen használhatók kép- és

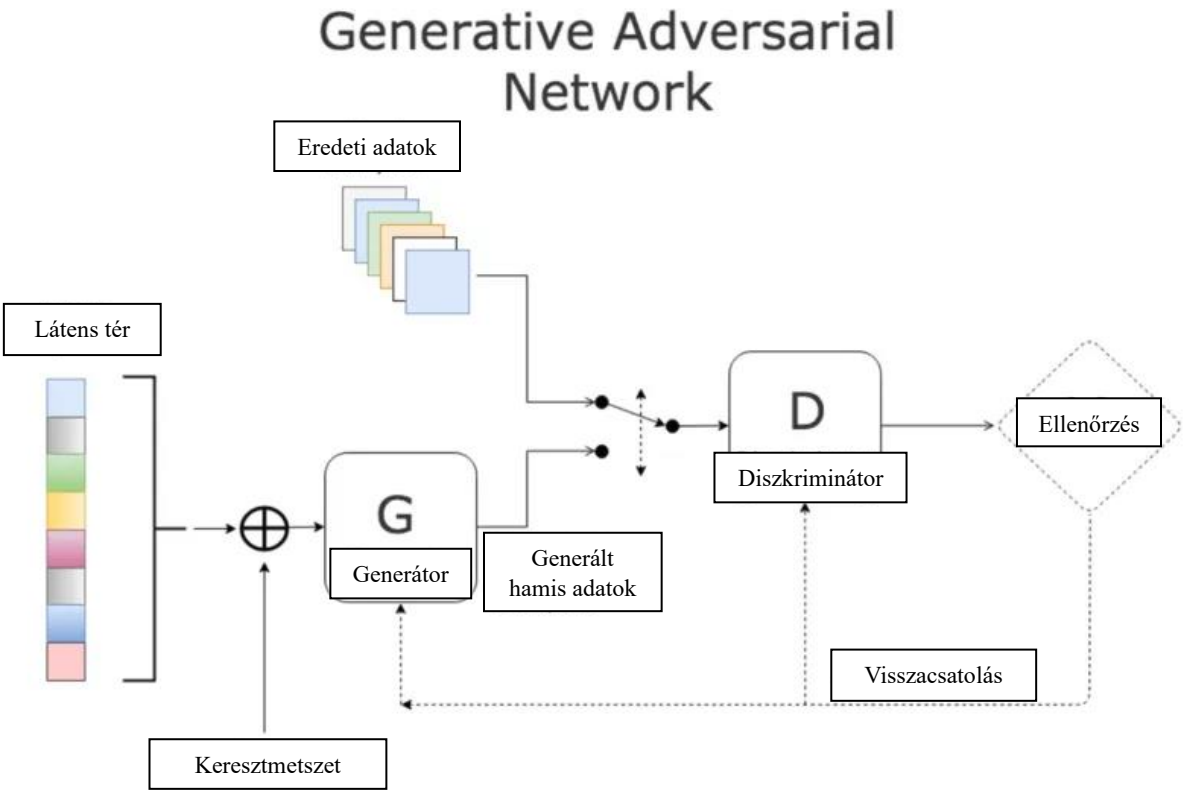
[3] Bart van der Sloot – Yvette Wagensveld – Bert-Jaap Koops: *Deepfakes: The legal challenges of a synthetic society,* Tilburg University, Tilburg, 2021, 3.
[4] A folyamat során a számítógépekbe matematikai adatmodelleket táplálnak be, közvetlen utasítás nélkül, amely által a gép az algoritmusok használatával azonosítja a mintákat, majd ezekkel létrehoz egy előrejelzésre alkalmas adatmodellt. Maga a rendszer alkalmazkodóképes, így gond nélkül használható abban az esetben is, ha az adatok folyamatosan változnak, a feladat típusa állandóan módosul vagy lehetetlennek tűnik egy adott kódolás.

hanganyagok gyártására, amelyek hasonlítanak a valódi tartalomhoz vagy éppen azokon alapulnak. Ezáltal egyfajta hamis valódiság ölt testet, amely beférkőzik a társadalmi-kulturális élet minden szegletébe, legyen az szórakoztatóipar, művészet, közösségi oldalak, politika vagy hírgyártás[5]. A legnagyobb fenyegetésnek azonban az igazságszolgáltatás van kitéve, ugyanis jelentős hatása lehet elsősorban a büntetőeljárásra, mivel lehet, hogy a (döntő) bizonyíték az MI segítségével gyártott hamisítvány. Ebből kifolyólag minden bizonyítékot fenntartással kell kezelni – még a hiteles és valós bizonyítékokat is –, amely igencsak befolyásolja az igazságszolgáltatás működőképességét, hiszen megrendül a jogrendszerbe vetett bizalom. Pedig a deepfake technológia fejlődése kétségbeejtően gyors, és továbbra sincsen egyértelmű jogi szabályozás, még a bizonyítási eljárás tekintetében sem, ami azt jelenti, hogy nem tudják teljes mértékben kiszűrni azokat az elemeket, amelyek MI segítségével készültek vagy azáltal hamisítottak. Ennek nyilvánvaló oka, hogy a már meglévő jogi normák inadekvátok ezen bizonyítékok hitelességét nézve, hiszen a hatályos jogi szabályozás a büntetőeljárás terén már jóval azelőtt létrejött, mielőtt a deepfake-kel kapcsolatos visszaélések visszaszorításának szükségessége mielőbbi megoldást követelt volna. Éppen ezért válik határokon átívelő problémává: egy ország sem volt felkészülve teljes mértékben az MI ily gyors ütemű terjeszkedésére.

A hamis kép-és hanganyagok meglehetősen gyors terjedésének oka, hogy meglepően könnyű azokat létrehozni. Egyik, mondhatni klasszikus módja a „Generative Adversarial Network” (GAN), ahol két gépi hálózatot használnak egy szintetikus, hamis tartalom fejlesztésére, egy ellentétes folyamaton keresztül. Az első hálózat a generátor. A létrehozandó tartalom típusát ebbe táplálják be, hogy adaptálja az adott információ jellemzőit. A generátor ezután megkísérel olyan új példákat létrehozni az adatokból, amelyek az eredeti adatok jellemzőivel megegyeznek, legalábbis azok tulajdonságait hordozzák, vagyis elkészíti a szükséges anyagot. Ezeket a generált mintákat aztán a második hálózatba táplálják be, amely szintén „megtanulja” azonosítani az adott adattípus jellemzőit. Ez az ún. diszkriminátor, amely megpróbálja felismerni a hibákat a példákban, és elutasítja azokat, amelyekről megállapítja, hogy nem az eredeti adatokkal megegyező tulajdonságokkal bírnak, azaz hamisítványként azonosítja őket[6]. Ezen hamisítványok aztán visszakerülnek a generátorba, hogy új adatokat állítson elő. Ez az eljárás addig folytatódik, amíg a generátor olyan valótlan tartalmat nem állít elő, amelyet a diszkriminátor valódiként nem azonosít. A GAN rendszerét először 2014-ben

[5] Delfino, Rebecca A.: *Deepfakes on Trial: A Call To Expand the Trial Judge's Gatekeeping Role To Protect Legal Proceedings from Technological Fakery*, Hastings Law Journal, San Fransisco, 2023, 293.
[6] Davis, Rama – Wiggins, Chris – Donovan, Joan: *Deepfake,* Belfer Center, Cambridge, 2020.

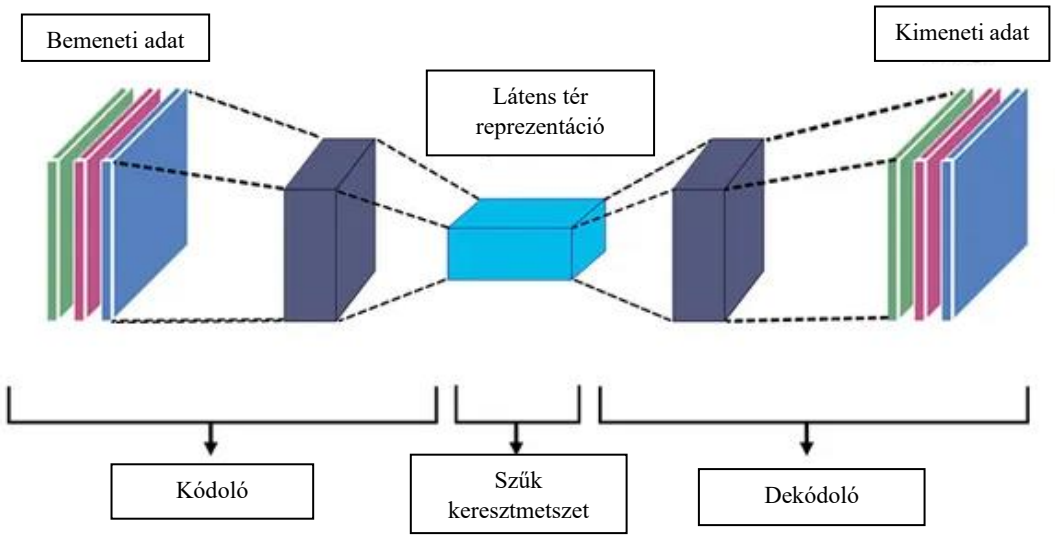
alkalmazták a gyakorlatban Ian Goodfellow és munkatársai, amikor bemutatták, hogy képesek emberi arcképek szintetikus másolatait megalkotni. Minél részletesebb a tartalom, amit a GAN hálózataiba táplálnak be, annál valósághűbb lesz a kimenet, vagyis a létrehozott kép.[7]



A GAN felépítése

Másik népszerű módja a deepfake tartalmak készítésének az AutoEncoders (AE), amely szintén két hálózatból áll: kódolóból és dekódolóból. A kódoló megtanulja hogyan tömörítse a betáplált adatokat egy kódolt formába, amely reprezentációt tartalmazó réteget nevezik szűk keresztmetszetnek. Ez a lehető legkisebb bemeneti adatdimenzió. Ha egy modell megtanulja, hogyan rekonstruáljon adatokat, és az a lehető legközelebb áll az eredeti bemeneti adathoz, akkor dekódolásról beszélünk. A két folyamat között lezajlódó jelenség a rekonstrukciós veszteség, amely egy módszer annak a meghatározására, hogy a dekóder milyen hatékonysággal működik és a kimenet milyen mértékben felel meg az eredeti bemenetnek.[8]

[7] https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf
[8] https://medium.com/@divakar1591/autoencoders-6fab1a9a9f9c



Az Autoencoders felépítése

A különböző alkalmazások leggyakrabban ezen módszereket használják, ily módon lehet minden nehézség nélkül arccserét, ajakszinkront, ún. motion transfert („mozgásátvitel”), illetve hang, kép és szöveg generálását végrehajtani.

2.2. Történeti áttekintés

A kép- és hanganyag hamisítás azonban nem új keletű dolog, ugyanis az első hamisítvány 1840-ben jelent meg, csupán két évvel azután, hogy Louis Daguerre elkészítette az első dagerrotípiát. A fényképet Hippolyte Bayard francia fotográfus szerkesztette, aki Daguerre riválisa volt a szakmán belül. Bayardot igen nagy veszteségek érték, hiszen áttörő sikert ért el, miután kifejlesztette a közvetlen képnyomtatás metódusát, ám érvényesítésével elkésett, mivel a Francia Tudományos Akadémia Daguerre fotografikai eljárását kezdte támogatni. Az őt ért igazságtalanságok miatt készítette el „Önarckép vízbefúltan” című fényképét, amelyen saját öngyilkosságát ábrázolja.[9] Bayard természetesen nem oltotta ki életét, de felhívta a figyelmet találmányára.

A 20. század viharos eseményei által viszont új szintre emelkedett a kép- és hanganyag hamisítás, valamint az álhírek terjedése, befolyásolva ezzel a politikai és gazdasági eseményeket. Jó példa erre az Amerikai Egyesült Államokban, 1950-ben kibontakozott Tydings-ügy, amikor Millard Tydings marylandi szenátor csalásnak, átverésnek és megtévesztésnek nevezte, hogy több száz kommunista dolgozott a külügyminisztériumban a

[9] https://petapixel.com/2012/11/15/the-first-hoax-photograph-ever-shot/

magát kommunistaellenesnek valló McCarthy szenátor hivatali ideje alatt[10]. Erre válaszul McCarthy hivatalnokai egy olyan képet szerkesztettek, amelyen Tydings lábható Earl Browderrel, az Amerikai Kommunista Párt vezetőjével, annak ellenére, hogy a két személy 1950 júliusáig soha nem is találkozott. A felvétel maga két fotóból volt összevágva: az egyik 1938-ban készült Tydingsról, ahogy éppen rádiót hallgat, a másik pedig 1940-ben Browderről, amint beszédet mond:

Bal és jobb oldalon az eredeti kép, középen az összevágott változat

A hamisítvány igencsak megingatta a szavazók Tydingsba vetett bizalmát és ennek következtében Tydings elvesztette a szenátusi választást John Butlerrel szemben.[11]

Ehhez hasonló, politikai hatással bíró szenzációnak számított a 2004-ben John Kerry amerikai külügyminiszterről elterjedt fotó, amin az 1970-es évek elején Jane Fonda színésznővel – aki aktivista tevékenysége miatt eléggé megosztó személyiségnek számított – tart békegyűlést, kifejezve ezzel, hogy elítélik a vietnámi háborút:

[10] Feeney, Matthew: *Deepfake Laws Risk Creating More Problems Than They Solve*, Regulatory Transparency Project of the Federalist Society, 2021.
[11] https://hoaxes.org/photo_database/image/the_tydings_affair

A fénykép hamar elterjedt, és az újságokon és a televízión keresztül bejárta a világot. Készítője, Ken Light azonban az évezred elején megkérdőjelezte a fotó valódiságát, és állította, hogy Jane Fonda nem is volt jelen az eseményen. Mint kiderült, a kép valóban hamisítvány volt és két felvételből állították össze: John Kerry 1971-es Long Island-i háborúellenes demonstrációját megörökítő és Jane Fonda 1972 augusztusában, a háborúellenes tüntetésen tartott beszédének pillanatából.[12]

Választási kampánnyal összefüggő hanganyag hamisítás volt az ún. Thatchergate, amelyet a brit anarcho-punk zenei együttes, a Crass hozott létre az 1983-as választás előtt. Margaret Tatcher és Ronald Reagan beszédeinek részleteit összevágva egy olyan hamis telefonbeszélgetést alkottak meg, amely tele volt politikailag káros részletekkel, aggodalmat okozva ezzel megannyi embernek.[13] A produktum ugyan cheapfake[14] volt, de a kiélezett nemzetközi helyzet idején, éppen, amikor a Falkland-szigeteki háború folyt, egyáltalán nem vetett jó fényt Thatcherre.

2.3. Deepfake alkalmazások és a mögöttük rejlő pszichológia

A valódi veszélyt mégsem a cheapfake, sokkal inkább a deepfake jelenti, amely technológia 2013-ban, maga a kifejezés pedig 2017-ben jelent meg a Redditen a „deep learning” (mélytanulás[15]) és a „fake” (hamis) szavak összevonásával.[16] Az „olcsó hamisítással”

[12] https://www.theguardian.com/media/2004/feb/17/newmedia.pressandpublishing
[13] https://www.aspi.org.au/report/weponised-deep-fakes
[14] A cheapfake vagy shallowfake olyan módosított médiatartalom, amelyet hagyományos és megfizethető technológiával változtattak meg, vagyis nem a gépi tanuláson alapulnak. Az olcsó hamisítási eljárások közé tartozik pl. a közösségi média felületein végzett photoshop (beleértve az arccserét), valamint a videók lassítása és gyorsítása.
[15] A mélytanulás olyan gépi tanulás, amely mesterséges neurális hálózatok által teszi lehetővé, hogy a digitális rendszerek tanuljanak és döntéseket hozzanak strukturálatlan, címkézetlen adatok alapján. A gépi tanulás az MI-rendszereket tanítja be úgy, hogy azok tanulni tudjanak a betáplált adatokból szerzett tapasztalatokból, illetve, hogy fel tudják ismerni a mintákat, javaslatokat tegyenek és alkalmazkodjanak az újonnan létrejövő struktúrákhoz.
[16] https://www.belfercenter.org/sites/default/files/2020-10/tappfactsheets/Deepfakes.pdf

ellentétben a mélyhamisítás egy olyasfajta technológia, ami gyakorlatilag önfejlesztő a mélytanulásra támaszkodva, ráadásul mindenki számára jutányosan elérhető, és – amint már említésre került – rendkívül nehéz felgöngyölíteni egyes eseteket.[17] Az interneten való rohamos terjedése egy reddit-felhasználóhoz köthető[18], aki „Deepfakes" néven töltött fel olyan videókat az oldalra, amelyekben híres színésznők arcát (pl. Gal Gadot, Taylor Swift) montírozta pornószínészek arcára „face-swap" technikával, realisztikus pornográf tartalmakat létrehozva ezzel.[19] Ezt követően egy másik anonim reddit-felhasználó megalkotta és szabadalmaztatta a FakeApp applikációt, amely legtöbb esetben ingyenesen teszi lehetővé felhasználóinak, hogy nehézség nélkül készíthessenek deepfake-et, ami eddig csak költségek árán és bonyolult technikai lépések útján volt lehetséges az arra képesített producereknek[20]. Természetesen nem ez az egyetlen, ún. nyílt forráskódú szoftver[21], ami gépi tanuláson alapul és bárki számára hozzáférhető[22]. Ilyen alkalmazás a Lyrebird, amely képes hangklónozásra is, és olyan hamis hangfelvételeket kreáltak vele, mint Trump elnök beszéde, ahol kifejti, hogy szankciókra van szükség Észak-Koreával szemben.[23] Ehhez hasonló tartalmak alkothatók a Face Shifter, a Deep FaceLab, az MSQRD, a Snapchat, a TikTok és még megannyi applikáció segítségével. Nem meglepő tehát, hogy olyan videók jelennek meg tömkelegével, amikről már nem lehet megállapítani teljes határozottsággal, hogy valósak-e vagy sem. A hamisított képek és hanganyagok töretlen terjedésének hátterében azonban nem csak ez az ok húzódik. Ezen okok közül az egyik, hogy több anyag kerül ki az internetre mindenkiről: fényképek, videó- és hangfelvételek, amelyek mind a deepfake-ek gyártását segítik. A kreált tartalmak terjedése továbbá azzal is indokolható, hogy az emberek leginkább azt tartják hitelesnek, amit saját szemükkel látnak. Ugyanis amikor valaki szemtanúja egy bizonyos eseménynek, történésnek, onnantól kezdve annak létezése, megtörténte általa kétségbe nem vonható és tényként kezelendő. Ez a jelenség az ún. „seeing is believing" frekvenciaillúzió, egy olyan kognitív torzítás, ami az észlelésből adódik. Röviden összefoglalva egy olyan heurisztikus folyamat, amely a problémamegoldásban és a döntési procedúrában segít, azonban nem tökéletes, hiszen

[17] Delfino, Rebecca A.: i.m. – 299. o.
[18] Meskys, Edvinas - Liaudanskas, Aidas - Kalpokiene, Julija - Jurcys, Paulius*: Regulating deep fakes: legal and ethical considerations*, Journal of Intellectual Property Law & Practice, 2020/15.
[19] https://www.britannica.com/technology/deepfake
[20] Gerstner, Erik: *Face/Off: „DeepFake" Face Swaps and Privacy Laws*, Defence Counsel Journal, 2020/1.
[21] A szabad vagy nyílt forráskódú szoftverek szabadon, bárki által másolható, használható, terjeszthető és módosítható számítógépes programok, ellentétben a zárt forráskódú eszközökkel, amik esetén egyetlen személy vagy szervezet a tulajdonos, és csak ők módosíthatják a szoftvert, illetve jogvédettség alatt állnak.
[22] Goggin, Benjamin: *From porn to 'Game of Thrones': How deepfakes and realistic-looking fake videos hit it big,* Business Insider, 2019.
[23] Langa, Jack*: Deepfakes, real consequences: Crafting legalisation to combat threats posed by deepfakes*, Boston University Law Review, Boston, 2021, 765.

a hiányzó információdarabkákat az egyén individuális érzékelése által pótolja, egy egyéni igazságot, valóságot kreálva ezzel.[24] Ennek eredményként felléphet olyan szisztematikus elfogultság, amely által az egyének túlbecsülik saját észlelési képességüket, ezáltal könnyebben válnak a deepfake áldozataivá.[25] A legjelentősebb vizsgálat a jelenséggel kapcsolatban egy 2012-es kutatás, ahol 5269 résztvevőt kérdeztek meg három valós és öt kitalált politikai eseményről. Minden koholt történést egy-egy fotó kísért, amely az állítólagos eseményt ábrázolta. A résztvevők mintegy fele tévesen emlékezett arra, hogy a szóban forgó eset megtörtént, 27%-uk pedig még arra is emlékezett, hogy a hírekben látta az eseményeket. Kimutatható volt egyúttal, hogy a politikai orientáció befolyásolta a hamis emlékek kialakulását, ugyanis a konzervatívok nagyobb eséllyel idézték fel tévesen azt, hogy Barack Obama kezet fogott az iráni elnökkel, míg a liberálisok nagyobb valószínűséggel emlékeztek arra, hogy George W. Bush az egyik híres baseball játékossal nyaralt a Katrina hurrikán katasztrófája alatt. A kutatás tehát alátámasztotta a tényt, hogy az események sokkal könnyebben beivódnak az emlékezetbe, ha azok kongruensek az egyén már meglévő viselkedésével és értékrendjével, hiszen a hamis események elősegítik a felismerés sikerét és az ismerősség érzetét keltik, amelyek azonban nincsenek összhangban az információ forrásával.[26] Viszont pontosan ebből származnak a problémák: a politikai megosztottság, a közbiztonságba vetetett hit megrendülése, illetve a társadalmi feszültség, amelyeket a deepfake még inkább csak fokoz.

3. A mélyhamisítás kockázatai és az abból származó hátrányok

A mélyhamisítások elsőre ártalmatlannak tűnhetnek, és az emberek átlagosan a közösségi oldalak különféle filtereire valamint játékaira asszociálhatnak, valójában korántsem veszélytelenek. Komoly fegyvert jelentenek ugyanis az olyan bűnözők számára, akik számítógépes környezetben követik el bűncselekményeiket. A hang- és képgenerálás nagymértékben hozzájárul az adathalászat sikerességéhez. Jó példa erre az a 2019-es ügy, ahol az elkövetők egy brit energetikai cég vezérigazgatójának hangját használták fel, 220 ezer euróval megkárosítva a vállalatot. Bizonyított továbbá az is, hogy az MI képes megzavarni a

[24] https://www.dw.com/en/frequency-illusion-when-seeing-is-believing-not-fact/a-64838404
[25] Köbis, Nils C.–Doležalova, Barbora–Soraperra, Ivan*: Fooled twice: People cannot detect deepfakes but think they can*, iSciense, 2021, 9.
[26] Frenda, Steven J.–Knowles, Eric D.–Saletan, William–Loftus, Elizabeth F.*: False memories of fabricated political events,* Legal Studies Research Paper Series, University of California, 2012, 280.

biometrikus szkennereket, ideértve az arcfelismerő rendszereket. Ezáltal az arccsere és más vizuális alapú deepfake programok pedig egyre csak növelik a bosszúpornó statisztikai számait.

A cyber támadások gazdasági szabotázs táptalaját is képezik: 2013-ban az Associated Press, az Egyesült Államok legrégibb és leghíresebb hírügynöksége egy Twitter posztban azt állította, hogy Obama elnök egy robbanás során megsérült. A hír komoly zuhanást idézett elő az amerikai tőzsdepiacon, ám később kiderült, hogy az ügynökség oldalát a Syrian Electronic Army, egy szíriai hackercsoport törte fel. Természetesen nem csak politikai vetülete lehet az ilyesfajta deepfake termékeknek, mivel egy meggyőző, ehhez hasonló csalás rendkívül káros a társadalomnak, főként, ha bűnügyi műveletekkel párosul.[27] A legingoványosabb terület talán a kiberbiztonság, hiszen a kiberbűnözők számára megvan a lehetőség, hogy az MI-t fegyverként használják fel. A technológia fejlődésével nyilvánvalóan megjelentek olyan szoftverek, amelyek védelmet nyújtanak, de előfordulhatnak olyan esetek, amikor a kiberbűnözők valamilyen módon rést ütnek a pajzson, igen nagy károkat okozva ezzel akár magánszemélyeknek, akár vállalkozásoknak. Az MI-rendszer által generált chatbotok[28] ugyanis másodpercek alatt képesek tiltott kódot generálni, egyszerűvé téve a szakértelemmel rendelkező bűnelkövetők számára, hogy rosszindulatú szoftvertámadásokat indítsanak. A károkat okozó kódokon kívül egyes eszközök képesek megtévesztő videókat generálni, amelyek a népszerű szoftverek bizonyos verziójának letöltéséről szóló útmutatónak látszanak, viszont a tartalom megtekintése közben ártalmas programokat telepítenek a célpontok eszközeire. Ezentúl az MI-t segítségül hívva a kiberbűnözők a jelszófeltörés sikerességének arányát is növelhetik és rekordgyorsasággal fedhetik fel az áldozatok bizalmas adatait.[29]

3.1. Kihatás a demokratikus rendszerekre

Az MI jelentősen veszélyeztetheti a demokratikus rendszereket, ugyanis a mélyhamisítás lehetővé teszi, hogy az „információs hadviselésben” felmerülő költségeket csökkenjenek, valamint az abban résztvevő szereplői kört kiszélesedjen, azaz egyre több emberhez jusson el az adott tájékoztatás. Ekképp a deepfake eszközei (leginkább a szöveggenerátorok) félrevezethetik a tömegeket, amely különösen hozzáad az ún. orosz propagandamodell értékéhez, ami leginkább a dezinformáció mennyiségét és gyorsaságát helyezi előtérbe a

[27] https://www.aspi.org.au/report/weaponised-deep-fakes
[28] A chatbot olyan szoftveralkalmazás, ami emberi erőforrás bevonása nélkül képes például ügyfelekkel történő kommunikációra. A chatbot-rendszereket úgy tervezték, hogy meggyőzően szimulálják azt, ahogyan egy ember viselkedne beszélgetőpartnerként.
[29] https://riskhubgroup.co.uk/wp-content/uploads/2024/01/5-Ways-Cybercriminals-are-Weaponsing-AI-1.pdf

hitelességgel és következetességgel szemben, annak érdekében, hogy a célközönséget összezavarja, elbizonytalanítsa és megossza. A jelenlegi államok ugyan rendelkeznek elég erőforrással ahhoz, hogy összehangolt, széleskörű kampányokat folytassanak, ám a fejlett nem állami szereplők már hajlandóságot mutatnak arra, hogy stratégiai hatású információs kampányokat indítsanak, gyengítve ezzel a kormányok stabilitását, felszínre hozva az etnikai feszültségeket. A közösségi médián keresztül terjedő, érzelmekre ható megtévesztő tartalmak, mint kontextus nélküli képek és téves állítások, etnikai erőszakot és gyilkosságokat eredményeztek többek közt Mianmarban, Sri Lankán és Indiában. Márpedig a deepfake egyre elérhetőbbé válik minden szférában, amiből az következik, hogy a nem állami csoportok egyre aktívabbak lesznek az online propagandában, ami növeli a szélsőséges szervezetek terjeszkedésének a lehetőségét.[30]

A mélyhamisítással kapcsolatos aggodalmak gyakran olyan esetekre összpontosulnak, ahol a hamisított anyag elég jó minőségűnek bizonyul ahhoz, hogy azt a szakértők kiszűrjék. Az ilyen típusú deepfake tartalmak megváltoztathatják egy választás eredményét, egy parlamenti vagy jogi eljárás menetét vagy akár egy diplomáciai, esetleg katonai törekvést. Ez azonban nem valószínű, hogy tömegesen bekövetkezhet, mivel a tájékozott és tapasztalt szakértő közönség feltehetően használja a rendelkezésre álló, felderítésre szolgáló eszközöket, hiteles bizonyítékokat keres, amelyeket a forrásuk és kontextusuk fényében értékel, valamint mérlegel, mielőtt a tartalom alapján cselekedne. El kell ismerni azonban, hogy vannak olyan szélsőséges esetek, amelyeknél a hiperrealisztikus deepfake-nek komoly hatása lehet: olyan szituációk, amelyekben az idő nagy szerepet játszik, és a tét hatalmas (nemzetközi válságok vagy váratlan, rendkívüli katonai helyzetek). A hamisított audiovizuális tartalmak ugyanis felhasználhatók a katonai parancsnokok helyzetfelismerésének rontására, politikai válsághelyzetben pedig bármely szereplő felhasználhat mélyhamisított tartalmakat az erőszak szítására.[31]

3.2. Az egyéneket érintő károk

A káros hatások ezen felsorolással még nem merülnek ki: a pletykák és hazugságok a társadalom létrejötte óta kísérik az emberiséget különböző mértékekben és formákban. Vannak közöttük olyanok, amelyek csupán irritálnak és zavarba hoznak, míg vannak, amelyek megaláznak és pusztítanak, sőt egyeseket erőszakra sarkallnak. Még inkább vonatkozik ez a deepfake-re, hiszen a benne rejlő hitelesség látszata miatt erőteljes mechanizmusként

[30] https://www.aspi.org.au/report/weaponised-deep-fakes
[31] https://www.aspi.org.au/report/weaponised-deep-fakes

szolgálnak egyeseknek arra, hogy kihasználjanak és szabotáljanak másokat. E tevékenység skálája sokrétű és taxatíve talán felsorolhatatlanok azon veszteségek, amelyek által keletkezhetnek. A kizsákmányolás megragadható a személyes adattal való (akár gazdasági célból való) visszaélés terén, de nem minden ilyen cselekmény ölt kézzelfogható formát. Vannak, amelyek más méltóságát degradálják le teljesen, jelentős erkölcsi károkat okozva ezzel, mivelhogy a deepfake technológiával az egyén szexuális identitását mások kielégítésére is felhasználhatják. Minden valószínűség szerint a mélyhamisított pornográf tartalmak áldozatainak többsége nő, és ne feledjük, hogy az internetes zaklatás és a bosszúpornó esetén is ez a tendencia, ami rávilágít egy fontos tényre: az MI által végrehajtott kihasználások nemi dimenziójára. Mary Ann Franks nyilatkozta ezzel kapcsolatban: „Ha te lennél a világ legrosszabb nőgyűlölője, ez a technológia lehetővé tenné számodra, hogy bármit elérj, amit csak akarsz."[32]

Az áldozatokkal kapcsolatos közvetlen pszichológiai károkkal összefüggésben említést kell tenni a jóhírnév és a becsület megsértéséről. Az élet minden területén – munkahelyen, párkapcsolatban, sportban, gazdaságban, politikában – az emberek képesek lesznek jelentős csapások mérni riválisaikra a deepfake technológia segítségével. A körülmények, az időzítés és a hamisítvány terjedésének függvényében a hatások elsöprőek és lesújtóak lehetnek. Bizonyos esetekben a hamisítvány leleplezése már túl késő lehet ahhoz, hogy a már bekövetkezett káros hatásokat orvosolni lehessen: gondoljunk például arra, hogy egy profi sportoló riválisa épp a draft[33] kezdetekor egy kompromittáló deepfake videót tesz közzé az illetőről, derékba törve ezzel karrierjét. Még ha később ki is derül, hogy a videóanyag hamisított, a következményeket (amelyek akár dollármilliók elvesztésével is járhatnak) lehetetlen visszafordítani, hiszen az elővigyázatos csapatok ezentúl inkább másokat választanak be, így az áldozat csak a draft későbbi fordulóiban nyerhet kiválasztást vagy teljesen ki is eshet[34]. Tény, hogy napjaink kommunikációs környezete fokozza az MI által okozott becsületsértés eseteinek számát, a kognitív torzítások és az algoritmus által terjesztett tartalmak kombinációja pedig növeli az ilyesfajta hamisítványok terjedésének esélyét. Amilyen könnyű az adatokat online másolni és tárolni, annál nehezebb a deepfake-ek eltávolítása, főleg, ha már egyszer közétették vagy megosztották azokat. Ezen tényezők a folyamatosan fejlődő keresőmotorokkal egyre csak

[32] Citron, Danielle K. – Chesney, Robert: *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security,* Berkeley School of Law, University of California, 2019, 1773.
[33] Angol kifejezés, a sportéletben a játékoskiválasztásra alkalmazzák, amely során a játékosokat a professzionális csapatok úgy választják ki (igazolják át), hogy a profiliga valamennyi csapata egyidejűleg juthat új játékoshoz.
[34] https://www.si.com/nfl/2016/04/29/laremy-tunsil-draft-controversy-ole-miss-dolphins-twitter-instagram-hacked

növelik annak veszélyét, hogy a munkáltatókhoz vagy az üzlettársakhoz is eljusson a hamis anyag, ami alapján még a hétköznapi üzleti lehetőségek is meghiúsulhatnak.[35]

3.3. A közvélemény befolyásolása és a politikai folyamatok torzítása

A deepfake nemcsak konkrét személyek vagy szervezetek számára jelent fenyegetést: számtalan módon képes kárt okozni a társadalomnak. Gondoljunk csak olyan esetekre, ahol hamisított kép- vagy videóanyagon egyes köztisztviselőknek rasszista megnyilvánulásai jelennek meg vagy éppen az, hogy megvesztegetik őket. A hamis videókon a politikai élet szereplői kémekkel vagy bűnözőkkel való találkozón is részt vehetnek, ami közfelháborodást válthat ki vagy éppen nyomozást indíthat el, teljesen alaptalanul. Ahogyan ezek a példák is mutatják, a mélyhamisítás veszélye rendszerszintű, így a károk torzíthatják a politikai kérdésekről folytatott demokratikus tárgyalásokat, manipulálhatják a választásokat, a jelentős köz- és magánintézményekbe vetett bizalmat áláshatják, megzavarhatják a katonai vagy hírszerzési műveleteket, hátrányt okozhatnak a nemzetközi kapcsolatok terén, mindezt a társadalmi megosztottság fokozásával és kihasználásával.

A közvélemény által, a szakpolitikai kérdésekről folytatott diskurzusok során igen sok hamis információ cserél gazdát, és bizony megeshet, hogy ezeknek a célja a vitában résztvevők lejáratása, hitelességük aláásása. A deepfake azonban még inkább súlyosítja ezt a helyzetet, mivel drámai módon növeli az álhírek terjedésének esélyét, amit számos politikai szereplő kihasználhat érdekei érvényesítése végett. Ahogyan azt a korábban említett orosz propagandamodell is mutatja, leginkább az állami szereplőknek van érdekeltségük az információk ferdítésében. Ám a demokratikus folyamatok leginkább akkor működőképesek, ha a politikai viták és igaznak vélt állítások tapasztalati érveken, bizonyítékokon alapulnak, és éppen ez az, ami a problémát jelenti: a deepfake-nek és az álhíreknek köszönhetően nincs általánosan elfogadott igazság, a vélemények megoszlanak, és az adott tények percről-percre változnak a gyors híráramlásnak köszönhetően. A valóság torzításával pedig a nemzeti és globális problémák megoldására lényegesnek és elsőrendűnek tűnő, de felesleges kérdésekbe bonyolódnak, amelyek a nagyobb gondokhoz képest eltörpülnek. Mindemellett ez a deformálódott valóság hatással van az egyének realitásérzékére: saját szubjektív valóságukban élnek, ahol hiedelmeiket kreált tényekkel támasztják alá. Az empirikus bizonyítékok tehát

[35] Citron – Chesney: i.m. – 1774.

érvényüket vesztik és az objektív felismerések is heves vitákat válthatnak ki, amely által a demokratikus folyamatok is nehezen haladnak tovább.[36]

A mélyhamisítványok terjedésének negatív hatásait az újságírók is érzékelik, ugyanis egyre inkább azzal a dilemmával szembesülnek, hogy vajon egy hírértékű eseményről szolgáltatott hang- vagy videóanyag hiteles-e? A hamis információk rendkívüli gyors terjedésének következtében ezt egyre nehezebb lesz megállapítani, amely által könnyűszerrel csapdába lehet csalni egy-egy hírszervezetet, és így a sajtó nem tud eleget tenni etikai és erkölcsi kötelességének[37], ami növelheti a társadalmi feszültséget és megosztottságot. A deepfake tehát elsősorban társadalmi aggályra ad okot, hiszen a deepfake tartalmak mind a jog, mind a demokratikus nyilvánosságot érintő kérdések területén rengeteg ellentmondást okoznak, így képesek negatívan befolyásolni a közvélemény alakulását.[38]

Azonfelül, hogy a deepfake tartalmak befolyásolják a politikai folyamatokat, képes a választás jelöltjeit is megalázni káros és hamis videó- vagy hangfelvételekkel, amely a szabotázs egy különösen kellemetlen módja. A választások kimenetelének csorbítására való lehetőség reális, főleg akkor, ha az elkövető úgy tudja időzíteni a hamisítvány terjedését, hogy ne legyen ideje a sértettnek azt megcáfolni. Ebből a szempontból vizsgálva a választás menete hasonló a korábban említett draft procedúrájához: mindkettő magában foglalja a döntésekhez szükséges ún. fojtópontokat, vagyis azon szűk időablakokat, amelyekben a visszavonhatatlan döntések születnek, és amelyekben a hamis információk által okozott károk helyrehozhatatlan következményekkel járhatnak. A 2017-es franciaországi választások jól szemléltetik ezt a helyzetet, ahol az oroszok egy olyan fedett akciót hajtottak végre a választások előtt kettő nappal (ez utólag a „Macron Leaks operation” nevet kapta), amely a kiberkémkedést és az információk manipulálását ötvözte annak érdekében, hogy megakadályozzák Emmanuel Macron elnökké választását. A terv részeként több mint húszezer digitális üzenetváltást és dokumentumot változtattak meg, majd hoztak nyilvánosságra. Rosszabb esetben olyan deepfake-et is elindíthattak volna, amely látszólag valós felvételekből áll, és olyan meggyőzően ábrázolja Macron elnököt, hogy a választás utolsó pillanatában elveszíti esélyét a francia szavazók bizalmára. Példának felhozható a 2024-es amerikai elnökválasztás is, amely összességében feszültséggel telibb, mint az eddigiek, és ebben a deepfake is igencsak jelentős szerepet játszik. Történt ugyanis, hogy Elon Musk, a Tesla vezérigazgatója egy manipulált kampányvideót

[36] Citron – Chesney: i.m. – 1777.
[37] Citron – Chesney: i.m. – 1785.
[38] Gosztonyi, Gergely – Lendvai, Gergely*: Deepfake és dezinformáció*, Médiakutató, Budapest, 2024, 43.

osztott meg, ahol Kamala Harris Joe Biden elnököt szenilisnek nevezi, és kijelenti, hogy fogalma sincs arról, hogyan kell egy országot vezetni, majd hozzáteszi, hogy nőként és színesbőrűként ő a sokszínűség igazi megtestesítője. A videó eredetileg egy Chris Kohls nevű konzervatív podcasterhez köthető, aki paródiának bélyegezte a bejátszást, azonban azzal, hogy Musk ezt újraposztolta, több mint 130 millió megtekintést ért el. Nem csoda, hogy a választópolgárokban egyre nagyobb riadalmat kelt az MI-alapú politikai dezinformáció.[39]

A választásokkal összefüggésben a mélyhamisítások által kiváltott fenyegetések nem feltétlen csak államilag támogatott titkos műveletekre vonatkoznak. Ugyanígy élhetnek vele nem állami szereplők, illetve olyan egyének, akiket a politikai folyamatok manipulálása motivál. Ha az ilyen beavatkozásokat precízen hajtják végre és megfelelően időzítik, az esetek nagy részében törvénytelenné tehetik a választási kampányt, illetve a választás kimenetelét, amely a közbizalom megrendülését és a belső stabilitás megingását jelentheti.[40] A deepfake ugyanis aláássa a köz- és magánintézményekbe vetett széleskörű bizalmat, beleértve a választott és kinevezett tisztviselőket, bírákat, nyomozókat, törvényhozókat és a követeket. A folyamat maga pedig öngerjesztő: ott, ahol már léteznek az erős bizalmatlansági narratívák, a provokatív mélyhamisítások könnyen széleskörű közönségre találnak, és ez növeli a politikai megosztottságot, amelynek eredményeként egyesek képesek lesznek hazugságokat és álhíreket terjeszteni a más nézeteket valló oldalról.[41]

Ám nemcsak az intézményeket veszik célpontba a deepfake-kel: erős fegyverként szolgálhat a társadalmi megosztottság növelésére, ezáltal igencsak gyengíteni lehet egy állam rendszerét, rendjét, sőt bizonyos esetekben a hamisított videók vagy hangfelvételek olyan mértékű érzelmi töltetettel bíró társadalmi mozgósítást indíthatnak el, amire az írott szavak – vagy pusztán a kijelentések – nem lennének képesek.[42] Gondoljunk csak egy olyan helyzetre, ahol a rendőri szervek és a helyi közösség között faji alapú feszültségek vannak. Egy kellően lázító hatású deepfake ez esetben jelentős zavargásokat idézhetne elő olyanfajta immateriális károkat okozva[43], amiket az adott közösségnek szabályszerűen meg kellene fizetnie. Mindazonáltal, hogy a társadalmi megosztottság kiélezi az ellentéteket és felszínre hozza az erőszakot, a közbiztonságot is jelentősen fenyegeti. A deepfake tartalmak vírusként terjednek, és ha az általuk támasztott hiperrealisztikus bizonyítékok meggyőzőereje önmagában nem

[39] https://www.france24.com/en/live-news/20240729-musk-faces-criticism-over-deepfake-kamala-harris-video
[40] Citron – Chesney: i.m. – 1778.
[41] Citron – Chesney: i.m. – 1779.
[42] Citron – Chesney: i.m. – 1780.
[43] Citron – Chesney: i.m. – 1781.

lenne elég, a közösségi média még inkább fokozza sokasodásukat. Mindez potenciális zavargásokhoz és pánikhangulathoz vezethet, amelyek következménye legjobb esetben gazdasági károk és egyéb, közhangulatot befolyásoló tényezők lehetnek, rosszabb esetben más tulajdonának rongálásával, személyi sérülésekkel és/vagy halált okozó tevékenységekkel járhatnak. Jó példa erre a 2018 elején történt pánikot keltő esemény, amikor a hawaii katasztrófavédelmi szerv egyik alkalmazottja lakossági figyelmeztetést adott ki egy közeledő ballisztikus rakétáról[44]. Kevésbé ismert, de céltudatos kísérlet volt a zavargások előidézésére az Orosz Internetkutató Ügynökség által indított kampány, amelynek keretében vegyi katasztrófa látszatát kívánták kelteni Louisianában és Ebola-járványt Atlantában. Mindkét esetben valós, de korlátozott mértékű kár keletkezett, bár hozzá kell tenni, hogy a történtekről szóló hírek nem terjedtek messzire bizonyítékok híján és a történtekkel összefüggésben felmerülő tényeket könnyedén ellenőrizhették.

Ám a deepfake terjedésével egyre kevésbé lesznek ennyire szerencsés kimenetelűek a pánik indukálására irányuló szándékos és rosszindulatú kísérletek. A mélyhamisítások ugyanis különösen hasznosnak bizonyulnak majd az ilyesfajta dezinformációs mozgalmak számára, növelve azok hitelességét. Képzeljük csak el, ha az atlantai Ebola-sztori mögött egy realisztikus hangfelvétel állt volna egy látszólagos telefonbeszélgetést rögzítve a Járványügyi Központ vezetőjével, aki kétségbeejtő híreket közöl, miközben segítséget kér az eset leplezésére a köznyugalom megőrzése érdekében.[45] Valószínűsíthető, hogy olyan szintű zúgolódás tört volna ki, amely köznyugalom elleni bűncselekmények melegágyaként szolgált volna.

3.4. Nemzetközi kapcsolatok stabilitásának megingatása

A mélyhamisítások a diplomáciai kapcsolatokat és a nemzetközi viszonyokat is megzavarják, főként akkor, ha a hamisítványt a közvélemény köreiben terjesztik. Nem minden állam rendelkezik ugyanis alkalmas intézményekkel ahhoz, hogy a diplomáciai stabilitást veszélyeztető deepfake tartalmakat kiszűrjék. A hamisítvány tehát akkor képes biztosan egy nemzetközi incidens kiváltására, ha egy vagy több államban a közvéleménnyel elhitetik, hogy valóban történt vagy elhangzott valami botrányos, korlátozva ezáltal az érintett kormány mozgásterét. A döntési fojtópont jelenségét tekintve egy jól időzített, a közvélemény hergelésére szánt deepfake terjesztése politikailag ellehetetlenítheti a célkeresztben álló

[44] https://www.npr.org/sections/thetwo-way/2018/01/30/581853255/hawaii-missile-drill-stated-this-is-not-a-drill-resulting-in-false-alert
[45] Citron – Chesney: i.m. – 1782.

résztvevőket.[46] Ilyesfajta, a közbiztonságot veszélyeztető és a nemzetközi kapcsolatok megzavarására szolgáló tevékenységek így a nemzetbiztonság megsértésének számítanak. Kiemelendőek ebben a tárgykörben még a katonai események, különösen a harci műveletek, ugyanis a deepfake jelentősen befolyásolja a stratégiai fontosságú, operatív vagy éppen taktikai bevetéseket. A dezinformációk terjesztésével a harcoló felek kölcsönösen törekszenek egymás megtévesztésére, és ez önmagában is hatalmas károkat és káoszt okozhat, kiváltképp, ha a civil lakosságot is érinti, amelynek körében erőszakos reakciókat válthat ki. Ide sorolhatjuk továbbá a kiber-kémkedést, amelynek révén az ellenfél valódi és bizalmas dokumentumokat szivárogtathat ki és hozhat nyilvánosságra.

A deepfake által okozott nemzetbiztonsági károk listája azonban a végtelenségig folytatható lenne, hiszen egyrészt teljesen eltérő esetekkel találkozhatunk, ebből kifolyólag pedig minden helyzetben más lehet a krízist kiváltó faktor; másrészt nincs egységes definíció a nemzetbiztonság meghatározására, így az világszerte egy szubjektív kategória.

4. A deepfake pozitív oldalai és a felmerülő szerzői jogok kérdése

4.1. A deepfake előnyei

Ahogy mindennek két éle van, az MI-nek sem csak és kizárólag hátrányai léteznek. Megemlítendő az az érték, amelyet az oktatás ügyében képes szolgáltatni: lehetőséget teremt a pedagógusok számára, hogy hagyományos tanítási eszközökön (pl. előadások, olvasmányok) túl érdekfeszítőbb és meggyőzőbb tanulási módszereket kínáljanak a diákok számára. Ez hasonló lépés a korábbi oktatási innovációhoz, amikor a különböző közösségi platformok megjelenésével könnyedén elérhetővé váltak az elektronikus tananyagok. A deepfake lehetővé teszi, hogy olyan tartalmakat állítsanak elő, amelyeken életre keltik a történelmi személyiségeket. A módosított tartalmak létrehozása azonban érdekes kérdéseket vet fel: vajon vonatkozik-e rájuk a szellemi jogok védelme és a szerzői jog? Lehet-e például egy háborús film jelenetét módosítani úgy, hogy egy parancsnok és jogi tanácsadója a háború következményeiről beszélgetnek, holott az eredeti jelenetben ez nem hangzik el?[47] Erre a dolgozat egy későbbi pontján kívánok kitérni.

A művészet terén kifejtett előnyös hatások valamilyen szinten kapcsolódnak az oktatáshoz is, bár nem feltétlen kell, hogy oktatási célt szolgáljanak. A meglévő technikáknak

[46] Citron – Chesney: i.m. – 1782.
[47] Citron – Chesney: i.m. – 1769.

köszönhetően elhunyt előadóművészeket támaszthatnak fel az elérni kívánt tömegek számára, amire jó példa a 2017-ben megjelent Star Wars: Az utolsó Jedik című film, ahol a rendezők úgy döntöttek, hogy az elhunyt színésznő, Carrie Fisher filmbeli dialógusait deepfake technológiával alkotják meg, a korábbi valós felvételeket felhasználva. Ebből is látszik, hogy nem messze van az a technikai váltás a filmiparban, amellyel a szinkronizált filmekben az adott színész a saját hangján szólalhat meg, bármelyik nyelven.

Ezentúl sok tartalomgyártó használja a mélyhamisítást közszereplők[48] és köztisztviselők parodizálására, kritizálására, valamint az aktivisták is alkalmazzák, hogy olyan módon demonstrálják álláspontjukat, amelyekre a szavak önmagukban nem lennének képesek.[49] Az aktivista tevékenységhez fűződő kampányok közül a legismertebb talán a „Malaria Must Die" kezdeményezés, ahol David Backham kilenc nyelven hívta fel a figyelmet a járvány elleni küzdelem szükségességére. A videón természetesen nem ő beszél, hanem szájmozgását teljesen az adott nyelven íródott szöveghez igazították.

A deepfake tartalmak ugyan nem tekinthetők származékos műveknek, de a legtöbbször szerzői jogi oltalom alá tartozó művek átdolgozásával jönnek létre, amely a szerzői jogról szóló 1999. évi LXXVI. törvény (továbbiakban: Szjt.) értelmében a szerző engedélyéhez kötött cselekmény. A deepfake előnyeire világítanak rá azok az esetek, amikor ismert, de már elhunyt zenészeknek vagy még élő énekesnek jelennek meg dalaik, amit az MI a korábbi dalok stílusát alapul véve hozott létre.[50] Ám ennek jogi és erkölcsi megítélése jelentős fejtörést okoz: az MI által előállított „termékek" nagy százalékban sértik a szerzői jogokat, azok közül is kiemelkedően a mű egységének védelmét[51], a név feltüntetéséhez való jogot[52], valamint az átdolgozáshoz való jogot.[53] A paródiákra ugyanakkor más szabályozás vonatkozik, mivel azok a szabad felhasználás körébe esnek. Az Szjt. kommentára alapján az új vonásoknak egyéni és egyúttal eredeti jelleget kell mutatniuk ahhoz, hogy paródiaműről legyen szó, azaz az eredeti

[48] Sorbán Kinga: *A bosszúpornó és deepfake pornográfia büntetőjogi fenyegetettségének szükségességéről,* Belügyi Szemle, 2020/10.
[49] Citron – Chesney: Citron – Chesney: i.m. – 1770.
[50] Herke Csongor: Deepfake: áldás vagy átok? Jogi szabályozási szempontok, Pro futuro, Debrecen, 2023, 162.
[51] Szjt. 13. § A szerző személyhez fűződő jogát sérti művének a becsületére vagy jóhírnevére sérelmes mindenfajta eltorzítása, megcsonkítása, megváltoztatása és a művel kapcsolatos más ilyen jellegű visszaélés.
[52] Szjt. 12. § (1) A szerzőt megilleti a jog, hogy művén és a művére vonatkozó közleményen – a közlemény terjedelmétől és jellegétől függően – szerzőként feltüntessék. A szerzőt a mű részletének átvétele, idézése vagy ismertetése esetén is meg kell jelölni. A szerző a neve feltüntetéséhez való jogot a felhasználás jellegétől függően, ahhoz igazodó módon gyakorolhatja.
[53] Szjt. 29. § A szerző kizárólagos joga, hogy a művét átdolgozza, illetve hogy erre másnak engedélyt adjon. Átdolgozás a mű fordítása, színpadi, zenei feldolgozása, filmre való átdolgozása, a filmalkotás átdolgozása és a mű minden más olyan megváltoztatása is, amelynek eredményeképpen az eredeti műből származó más mű jön létre.

mű szerzőjétől ne kelljen engedélyt kérni. A paródia tehát csak akkor egyéni és eredeti jellegű alkotás, ha az eredeti mű vonásai benne elhalványulnak és az új egyéni vonások kerülnek túlsúlyba.[54] A szerzői jogok védelme a deepfake vonatkozásában természetesen rögzítésre került Európai Uniós szinten is: az MI-ről szóló rendelet 50. cikk (4) bekezdése kimondja, hogy „amennyiben a tartalom nyilvánvalóan kreatív, szatirikus[55], művészi vagy fikciós filmművészeti alkotás, videójáték vizuális anyaga, vagy jogi szabályozási szempontok hasonló mű vagy program részét képezi, a deepfake-kel kapcsolatosan meghatározott átláthatósági kötelezettségek az ilyen létrehozott vagy manipulált tartalom meglétének megfelelő, egyértelmű és látható módon történő, a mű megjelenítését nem akadályozó közlésére, valamint adott esetben az alkalmazandó szerzői jogok közlésére korlátozódnak."

4.2. Az általános adatvédelem követelménye

A szerzői jogokhoz kapcsolódóan említést kell tenni az általános adatvédelemről, hiszen a deepfake tartalmaknak, főleg a videóknak személyesadat-jellegük van, mivel az abban szereplő „személy" soha nem tanúsította az adott magatartást, nem hagyták el a száját az adott mondatok. Ez esetben tehát egy olyan mesterségesen generált tartalomról szól, amely nem az értintettről készült közvetlenül és nem is tőle származik, ugyanis magán a videón nem ő szerepel, hanem egy szintetikus entitás. A kérdés az, hogy olyan elkülönült jogi sorsú-e a deepfake videó, ami nem minősül az utánzott személyes adatának, vagy inkább a testi és más külső jegyek nagyfokú hasonlósága miatt a tartalom az utánozni kívánt személy személyes adata lesz-e? A GDPR meghatározása szerint a személyes adat fogalmának kulcseleme, hogy a természetes személy azonosítható legyen, akár indirekt, akár direkt módon. Azonosíthatóvá egy személy a vele kapcsolatba hozható bármely információ, úgymint testi, fiziológiai vagy szociális azonosságára vonatkozó tényezők alapján válhat. Ez alapján elmondható, hogy nem feltétlenül szükséges egy természetes személy azonosíthatóságának megállapításához, hogy az információ közvetlenül tőle származzon, elég, ha rá lehet ismerni a róla közölt jellemzők alapján.[56] Ezért a deepfake videó is a megszemélyesített, mások által beazonosítható, egy adott csoport (vagy a társadalom) minden más tagjától elkülöníthető természetes személy adata lesz, így alkalmazandók rá a GDPR előírásai. Emiatt a tartalom nyilvánosságra hozatalára és kezelésére csak akkor

[54] https://copy21.com/2020/03/hiba-a-matrixban-nicolas-cage-neo-szerepeben-avagy-a-deepfake-ai-videok-szerzoi-jogi-szemmel-1-resz/
[55] Mráz szerint a kifejezetten szatirikus ábrázolásnak bele kell férnie a vélemény-nyilvánítás szabadságába. Mráz, Attila: *Deepfake, demokrácia, kampány, szólásszabadság*, In: Török Bernát – Ződi Zsolt (szerk.): *A mesterséges intelligencia szabályozási kihívásai,* Ludovika Egyetemi Kiadó, Budapest, 2021.
[56] Aczél Petra – Veszelszki Ágnes: *Deepfake: A valótlan valóság*, Gondolat Kiadó, Budapest, 2023, 147.

kerülhetne sor, ha jogszerű céllal és jogalappal rendelkezik az azt nyilvánosságra hozó személy, mivel ez a cselekmény a magánszférára gyakorolt hatása miatt alapvető kockázatokat rejt magában.[57]

4.3. Az önkifejezés szabadsága

A mélyhamisítás az önkifejezés szabadságát is egy új szintre növeli. Természetesen nem minden, az MI által önkifejezési céllal létrehozott alkotás művészeti jellegű: egyesek az „avatar élményért" veszik igénybe. A különböző videójátékok bizonyos mértékig lehetővé teszik a felhasználó számára, hogy olyan élményeket éljen át vagy érzékeljen, amelyek a valóságban egyébként lehetetlenek, veszélyesek lennének vagy következményük nem kívánt eredménnyel járna.[58] A nagy népszerűségnek örvendő Black Mirror sorozat epizódjainak többsége pontosan e téma köré építenek fel egy jövőbeli disztópikus világot. A „Be Right Back" című rész egy olyan szolgáltatást mutat be, amely azt ígéri az embereknek, hogy kommunikálhatnak elhunyt szeretteikkel. Az online platform az elhunyt hozzátartozó addigi üzenetváltásait és a közösségi térben végzett tevékenységeit dolgozza fel, majd segítségükkel készít egy új virtuális verziót a személyről. A szolgáltatáshoz kiegészítést is ajánlanak, egy androidot, amely tulajdonképpen a hozzátartozó szintetikus teste. A magányos, gyászoló főhősnő így él is a lehetőséggel, ám hamar kiderül, hogy a szituáció korántsem olyan, mintha az illető élne: a robot nem képes az alapvető emberi tevékenységekre, teljesen természetellenes, hiszen nem alszik, mindennek engedelmeskedik, nincsenek önálló gondolatai, vagyis az elhunytra jellemző tulajdonságokat nem érzékelteti. A főszereplő már a pszichés összeomlás szélén áll, amikor véget vet az egésznek és megsemmisíti az avatart az androiddal együtt. Másik kiemelendő epizód a sorozatból a „Joan is Awful" címet viseli. A fő karakter, Joan egy átlagos nő, aki rájön, hogy egy globális streamingszolgáltató napról-napra adaptálja életét, minden titkát és magánügyét kiteregetve, egyfajta reality show-t csinálva az életéből. Szerepét Salma Hayek színésznő játssza a „sorozatbeli filmsorozatban", illetve neki is csak az MI által generált változata. Úgy látszik, az őrületből nincs kiút, így Joan elvetemült lépésekre szánja el magát, és olyanokat művel, amelyeket Hayek a való életben sem tud tovább tétlenül nézni, és meg akarja szakítani a szereplését, mondván, hogy ő nem adja a nevét ehhez. Ám egyikőjük sem szabadul könnyen, hiszen a szerződésben oldalakon keresztül taglalja a szolgáltató, hogy bármit kiadhatnak Joanról, és az egyezség fel nem bontható. Ők ketten pedig aláírták a szerződést, így egyetlen

[57] Aczél – Veszelszki: i.m. – 148.
[58] Citron – Chesney: i.m. – 1770.

megoldás van: elpusztítani azt a szigorúan őrzött helyen lévő kvantum számítógépet, amely az adatokat másodpercről-másodperce dolgozza fel a valós életből és közvetíti eseténként. A sorozat többi része is rendkívül elgondolkodtató és egy olyan világképet vetít előre, amely a technológia csúcsának tűnik, ugyanakkor mérhetetlenül ijesztő.

Az önkifejezés területén azonban a legjelentősebb a fogyatékossággal élő emberek életének megkönnyítése, például a bénulás bizonyos formáiban szenvedő személyek (pl. ALS) beszédképességének valamilyen szintű helyreállítása, de ugyanez megvalósítható siketnémák esetében egyaránt. A deepfake technológia által lehetséges az is, hogy ha valaki élete során veszítette el hangját, akkor a korábbról fennmaradt hangfelvételek alapján saját hangján szólalhasson meg.[59]

5. A deepfake jelenlegi jogi szabályozása, különös tekintettel az EU szabályokra

Hazánkban nincs alapvető, átfogó, ezt a területet szabályozó jogi norma. A deepfake-kel kapcsolatban elkövetett bűncselekmények esetén kizárólag a Büntető Törvénykönyvről szóló 2012. évi C. törvény (a továbbiakban: Btk.) azon tényállásai relevánsak, amelyekkel az adott cselekmény összefüggésbe hozható, vagyis amelyeket a visszaélést megvalósító elkövetett.[60] Ilyen lehet a személyes adattal visszaélés (Btk. 219. §), a zaklatás (Btk. 222. §), a becsület csorbítására alkalmas hamis hang- vagy képfelvétel nyilvánosságra hozatala (Btk. 226/A. §), a rémhírterjesztés (Btk. 337. §), a szerzői vagy szerzői joghoz kapcsolódó jogok megsértése (Btk. 385. §) vagy a becsületsértés (Btk. 227. §). Utóbbi igen komoly kérdéseket vet fel a jogszabály harmadik bekezdésében[61] megfogalmazottak miatt. Nehéz ugyanis meghúzni a határt a sajtóban megjelenő kifigurázások, szatírák emberi méltóságot sértő és humoros volta között. Ilyesfajta dilemmát vetett fel a 2014-es országgyűlési választások során a Negyedik Köztársaság Párt által készített kampányvideó, amin „egy katonai egyenruhába bújt majomnak öltözött ember szerepel, aki a korábbi magyar miniszterelnökök hangjára tátog”, és amiről alkotmánybírósági határozat is készült.

[59] Herke Csongor: *Deepfake: áldás vagy átok? Jogi szabályozási szempontok*, Pro futuro, Debrecen, 2023, 165.
[60] Herke Csongor: i.m. – 174.
[61] Btk. 227. § (3) Nem büntetendő becsületsértés miatt annak a cselekménye, aki az (1) bekezdésben meghatározottakat a közügyek szabad megvitatása körében sajtótermék vagy médiaszolgáltatás útján követi el, feltéve, hogy cselekménye nem irányul a sértett emberi méltóságának nyilvánvaló és súlyosan becsmérlő tagadására.

A deepfake-kel összefüggésbe hozható cselekményekkel kapcsolatban az Európai Parlament által 2024. június 13-án elfogadott, az MI-re vonatkozó harmonizált szabályok megállapításáról szóló rendeletre támaszkodhatunk. Ennek preambuluma hangsúlyozza, hogy a norma célja nem más, mint az emberközpontú és megbízható MI elterjedésének előmozdítása, a belső piac működésének javítása, valamint Unió-szerte az egészség, a biztonság, az alapvető jogok, a demokrácia és a jogállamiság, illetve a környezet magas szintű védelme az MI-rendszerek[62] káros hatásaival szemben. Ezen célok rögzítése igencsak fontos, hiszen vitathatatlan tény, hogy a technológia befolyásolhatja a felsorolt tényezőket, mégpedig negatív módon. A rendelet megfogalmazásában a deepfake „manipulált vagy szintetikus hang, kép vagy videótartalom, amely megtévesztő módon az eredetiség vagy hitelesség látszatát kelti, és amely úgy ábrázol személyeket, mintha olyasmit mondanának vagy tennének, amit soha nem mondtak vagy tettek, és amelyet az MI-n alapuló technikákkal – beleértve a gépi tanulást és a mélytanulást is – állítanak elő". Ahogyan a becsületsértés tényállásának esetén, úgy az Európai Unió rendeletében is rögzítik, hogy az MI-rendszer felhasználása engedélyezett, amennyiben a létrehozott tartalom kreatív, művészi vagy filmművészeti alkotás, videojáték vizuális anyaga, illetve hasonló mű vagy program részét képezi, továbbá, ha azt bűnüldöző hatóságok használják fel a deepfake felderítésére, annak használatához kapcsolódó bűncselekmények megelőzésére, kivizsgálására és büntetőeljárás alá vonására. A rendelet tárgyi hatálya kiterjed az MI által vezérelt rendszerek Európai Unión belül forgalomba hozatalára és üzembe helyezésére is, valamint meghatározza az MI-gyakorlatra vonatkozó tilalmakat, és a nagy kockázattal bíró MI-rendszerekre vonatkozó különös követelményeket. Személyi hatályát tekintve kiterjed a forgalomba hozó vagy üzembe helyező szolgáltatókra, az MI EU-n belüli felhasználóira, illetve a rendszerek harmadik országban található szolgáltatóira és felhasználóira, ha az előállított kimenetet[63] az Unióban használják.[64]

A MI-rendelet az MI-rendszereket kockázatuk alapján osztja fel: tiltott, nagy kockázatú, korlátozott kockázatú, illetve minimális kockázatú vagy nem kockázatos rendszerekre. A III. melléklet nevesíti azon példákat, amelyeket mindenképp magas kockázatúnak kell tekinteni. A felsorolásban nem szerepelnek a deepfake tartalmak létrehozásához szükséges MI-rendszerek,

[62] MI-rendszer: olyan gépi alapú rendszer, amelyet úgy terveztek, hogy különböző szintű autonómiával működjön, és amely explicit vagy implicit célok érdekében olyan kimeneteket, például előrejelzéseket, ajánlásokat vagy döntéseket képes generálni, amelyek befolyásolják a fizikai vagy virtuális környezetet.
[63] A kimenet (output) a rendelet szerint a felhasználó által meghatározott célkitűzés alapján a MI-rendszer által adott eredmény (pl. generált tartalom, előrejelzés, ajánlás, döntés, meghatározott feltételek alapján összeállított szöveg vagy következtetés).
[64] Aczél – Veszelszki: i.m. – 158.

a bűnüldöző hatóságok által a felderítésre szolgáló rendszerek viszont igen. Éppen ez az oka annak, hogy az Európai Unió egységes piacán történő forgalomba hozatal előtt a magas kockázatú rendszereknek szigorú követelményeknek kell megfelelni. Ennek értelmében tehát a deepfake-ek létrehozásához szükséges struktúráknak – korlátozott kockázatú rendszerek lévén – nem kell a merev, ellentmondást nem tűrő feltételekhez igazodniuk, de természetesen meg kell felelniük a rendelet által előírt átláthatósági követelményeknek. Emiatt azon MI-rendszerek felhasználói, akik olyan meglévő személyekre, tárgyakra, helyekre vagy más szervezetekre, esetleg eseményekre érzékelhetően hasonlító képet, audio- vagy videótartalmat generálnak vagy manipulálnak, amely egy személy számára megtévesztő módon eredetinek vagy valódinak tűnhet, kötelesek azt közölni, illetőleg a tartalmat oly módon jelölni, amelyből megállapítható, hogy a tartalom mesterséges úton keletkezett vagy manipulálták azt. Ez a kikötés nem vonatkozik azokra az esetekre, amikor a felhasználást törvény engedélyezi, vagy ha az a véleménynyilvánítás szabadságához, továbbá a művészet és tudomány szabadságához való jog gyakorlásához szükséges, és a harmadik felek jogaira és szabadságaira vonatkozó megfelelő biztosítékok hatálya alá tartozik. Ezek alapján nem minősül szabálysértőnek, ha egy, az Európai Unión kívüli országban található felhasználó a művészet vagy a véleménynyilvánítás szabadságának keretében például ismert politikusok képmását felhasználva készíti el a tartalmat, ha azt megfelelő jelzéssel látja el. Az MI-rendelet nem határozza meg a konkrét jelölési módokat, sem a közlési forma követelményeit, aminek hátránya, hogy iránymutatás nélkül valószínűleg számos eltérő megközelítést fog eredményezni, és ez nem nyújt egységes megoldást.[65]

A rendelet megsértése és az MI-gyakorlatok tilalmának be nem tartása esetén többszintű szankciórendszert állapítottak meg a jogalkotók, amely legfeljebb 35.000 EUR összegű, vagy ha az elkövető vállalkozás, az előző pénzügyi év teljes világpiaci árbevételének legfeljebb 7%-át kitevő összegű közigazgatási bírsággal sújtandó, a kettő közül a magasabb összegűt figyelembe véve. A kis- és középvállalkozásokra és az induló innovatív vállalkozásokra pedig arányos mértékű közigazgatási bírságokat szabhatnak ki. Ha az MI-rendszer nem felel meg a gazdasági szereplőkkel vagy a bejelentett szereplőkkel kapcsolatos rendelkezések bármelyikének, az legfeljebb 15.000 EUR összegű, vagy ha az elkövető vállalkozás, az előző pénzügyi év globális éves árbevételének legfeljebb 3%-át kivető összegű közigazgatási bírsággal büntetendő, a kettő közül szintén a magasabb összegűt figyelembe véve. Ha pedig a bejelentett szervezetek vagy az illetékes nemzeti hatóságok kérésére helytelen, hiányos vagy

[65] Aczél – Veszelszki: i.m. – 159.

félrevezető információ szolgáltatására kerül sor, legfeljebb 7.500.000 EUR összegű, vagy ha az elkövető vállalkozás, az előző pénzügyi év teljes globális és éves árbevételének legfeljebb 1 %-át kitevő összegű közigazgatási bírsággal sújtandó, a kettő közül a magasabb összegűt figyelembe véve.[66]

A fentiek alapján megállapítható, hogy a felhasználó felelős az átláthatósági követelmények betatásért, azok megszegése esetén a felhasználóra hárul a felelősség, aki bírságolható. Ez azt is jelenti, hogy a nem megfelelően jelölt deepfake tartalmak miatt sem az azok létrehozását lehetővé tevő MI-rendszereket üzemeltető szolgálató, sem az importőr, sem pedig a forgalmazó nem vonható felelősségre. Nem tartozik továbbá a rendelet hatálya alá és nem szankcionálható az sem, aki a felhasználó által az MI-rendszer alkalmazásával létrehozott, a kikötéseket figyelmen kívül hagyva nem megfelelően vagy egyáltalán nem jelölt tartalmakat másokkal megosztja, azokat terjeszti. Kérdéses azonban, hogy a szankciórendszer az eleve ártó szándékkal megalkotott illegális tartalmak létrehozóival szemben mekkora visszatartó erőt jelent[67], illetve, hogy mennyire alkalmas az MI-vel összefüggésben elkövetett bűncselekmények visszaszorítására. Ennek fényében és az MI gyors fejlődésére tekintettel a 112. cikk (3) bekezdésében foglaltak alapján a Bizottságnak 2029. augusztus 2-ig, majd azt követően négy évente jelentést kell benyújtania az Európai Parlamentnek és a Tanácsnak a rendelet értékeléséről és felülvizsgálatáról. A jelentésnek tartalmaznia kell a végrehajtás struktúrájának értékelését, valamint azt, hogy szükség van-e valamely uniós ügynökségre a feltárt hiányosságok orvoslásához. A megállapítások alapján az említett jelentést adott esetben a rendelet módosítására irányuló javaslatnak kell kísérnie és azt közzé tenni.

Következésképpen elmondható, hogy bár uniós szinten volt előrelépés a deepfake szabályozását illetően és sikerült az MI fogalmát összekapcsolni a védendő társadalmi értékekkel, be kell látnunk, hogy amilyen kiterjedt a tárgykör, ezen törekvés még kevés ahhoz, hogy teljes hatékonysággal lehessen védekezni az MI által okozott károk ellen, akár a polgári jog, akár a büntetőjog szintjén.

[66] Az Európai Parlament és a Tanács 2024. június 13-i (eu) 2024/1689 rendelete a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet).
[67] Aczél – Veszelszki: i.m. – 160.

6. A deepfake hatása a büntető eljárásjogra

Az állami és nemzetközi törvények korlátozott száma és hatálya miatt a hamisítványok behatoltak a demokrácia, valamint az igazságszolgáltatás intézményeibe, azaz az egész jogrendszerbe. A deepfake bizonyítékok már jelenleg is kritikus pontjait képzik mind a polgári, mind a büntetőpereknek, és kiemelt figyelmet kapnak. Ennek ellenére a bizonytalanság mindig ott lebeg minden deepfake bizonyítékokat magában foglaló ügy felett. Az ebből eredő ártalmak ugyanis nem csak az igazságszolgáltatás működését torzíthatják súlyosan, de a sértettek számára is érzelmi kárként vagy jogi költségek formájában jelentkezhetnek.[68]

A jogrendszer azonban nem csak a hamis bizonyítékok által sérülhet: az is előfordulhat, hogy valaki a valós bizonyítékról állítja, hogy hamis, miközben tisztában van azzal, hogy valójában nem az, vagyis a legártalmasabb hazugságok tagadások formájában jönnek létre. Igaz, hogy önmagában véve nagy kockázatot rejt egy ilyesfajta stratégia, de kevésbé rizikós akkor, ha az ügyet nem övezi médianyilvánosság és a nyomozó hatóságok nem rendelkeznek a csalás leleplezéséhez szükséges technikai eszközökkel. A deepfake „hasznosnak bizonyul" olyan esetekben is, amikor a vádlott az igazság elől menekül. Ironikus módon a valódi tetteikért és szavaikért felelősséget nem vállaló személyek hitelesebbé válnak, ha felhívják a közvélemény figyelmét a mélyhamisítások jelentette veszélyekre[69]; amíg pedig nem lehet kétséget kizáróan a bizonyíték valódiságát alátámasztani, a hivatkozás érvényben maradhat és akár felmentéshez is vezethet[70]. Ezt a veszélyt a „hazugok osztalékának" nevezzük („The Liar's Divident").

A problémát azonban nem szabad túlértékelni sem, hiszen az elfogultság is diszfunkcionalitáshoz vezethet. Bizonyos, hogy a hamis bizonyítékok nem érintenek minden bírósági eljárást, illetve nem minden esetben kerülnek elő digitális vagy hangalapú bizonyítékok. Mindenesetre az igazságszolgáltatás jelenlegi legnagyobb megpróbáltatásai közé tartozik egy olyan megbízható, kifinomult keresőrendszer felépítése, amely kiszűri a manipulált bizonyítékokat, és megoldást talál a hitelességgel kapcsolatban felmerülő kérdésekre.[71]

[68] Delfino, Rebeca A.: i. m. – 305.
[69] Citron – Chesney: i.m. – 1786.
[70] Palmiotto, Francesca: *Detecting Deep Fake Evidence with Artificial Intelligence. A Critical Look from a Criminal Law Perspective*, SSRN, 2023.
[71] Delfino, Rebeca A.: i.m. – 307.

6.1. Kihívások a bíróságon: bizonyíték, védelem és bírói szkepticizmus

A deepfake három alapvető kihívást jelent a tárgyalásokon: annak bizonyítása, hogy az adott digitális kép- vagy hanganyag hiteles; mérlegelni azt az állítást, miszerint a bemutatott bizonyíték mélyhamisítás; illetve a bíróság növekvő bizalmatlanságának és kételyeinek eloszlatása a bemutatott anyagokkal kapcsolatban.

A deepfake bizonyítékok hitelességének megállapítása a hagyományos eszközökkel időigényesnek, költségesnek és némely esetben kivitelezhetetlennek mutatkozik. Ezen a technológia gyors fejlődése sem segít, ugyanis szabad szemmel előbb-utóbb képtelenek leszünk észrevenni azokat a kifinomult részleteket, amelyek arra utalnak, hogy az adott kép-, videó- vagy hanganyag hamisítvány. Persze megoldás lehetne, hogy a tanút kérjük fel a bizonyíték hitelesítésére, a mélyhamisítás azonban gyakran a valódi és hamisított anyagok keveréke, így egy tanú is képtelen lehet teljes mértékben felismerni a hibát és rekonstruálni a történteket.[72] Ezt befolyásolhatja továbbá a fentebb már említett kognitív torzítás jelensége is.

A büntetőeljárás során a védőknek nemcsak az okozhat nehézséget, hogy a bizonyítékok valódiságát alátámasszák, hanem az is, hogy a terhelt ellen felhozott, hamisítással alátámasztott érveket megcáfolják. A cheapfake hamisítványokat könnyű leleplezni, hiszen léteznek eredeti tartalmak, amelyekkel azokat össze lehet hasonlítani. A deepfake-ek feltárása viszont problémás lehet, mivel általában nincs eredeti tartalom, így nincs összehasonlítási alap sem. Másrészt, ha a védő a bizonyítási eljárásban hitelesíteni kívánt videóanyagról úgy ítéli meg, hogy annak bizonyítóerejét felülmúlják a hitelesítéssel járó költségek, és azokat a terhelt vállalja, akkor is benyújthatja a bizonyíték kizárása iránti indítványát, ám fennáll a veszélye annak, hogy a bíróság elé olyan hamis bizonyíték kerül, amely alapján a védelem felelősségre vonható. Ezen igazságszolgáltatásra gyakorolt hatás a büntetőügyekben különösképp érződik: a terhelt szabadsága (egyes országokban élete) múlhat azon, hogy a védelem fel tudja-e hozni a megfelelő érveket az ellene támasztott bizonyítékok hitelességének cáfolatára, illetőleg a vizsgálathoz szükséges forrásokat össze tudja-e gyűjteni. Ha egy hamisítványt minden kétség nélkül elfogad a bíróság, akkor egyúttal távolabb is kerülnek az igazságtól és a méltányosságtól. Ha a terheltnek van alibije vagy azt alátámasztó tanúja, aki hajlandó vallomást tenni, de nincs mód arra, hogy a bűncselekmény elkövetésének idején és helyszínén őt ábrázoló egyértelmű audio/vizuális/audiovizuális bizonyítékot megcáfolják, nehéz elképzelni, hogy a bíróság a tanúvallomást elfogadja. Természetesen a vád oldaláról is jelentkezhetnek problémák. Ha a

[72] Delfino, Rebecca A.: i.m. – 309.

meglévő eszközök nehézkesnek vagy költségesnek bizonyulnak a hitelesség ellenőrzése szempontjából, az ügyészség dönthet úgy, hogy nem emel vádat, még akkor sem, ha valószínűsíthető, hogy a digitális bizonyíték valódi.[73]

A hamisítványok továbbá lehetőséget teremtenek arra, hogy az eljárásban résztvevő felek újabb kérdéseket, érveket és kifogásokat vessenek fel a valós bizonyítékokkal szemben, így kétséget ébreszthetnek a bíróságban minden digitális hang és kép hitelességét tekintve, még akkor is, ha tudatában vannak a bizonyíték valódiságának. Ez az ún. mélyhamisítás-védelem[74], amely összefüggésben van az imént említett „hazugok osztalékával".

A pszichológia szerint az emberek a vizuális érzékelést előnyben részesítik mindenhez képest, így képesek azt az igazság mutatói fölé is helyezni. Amerikai kutatások eredményei alapján azok az esküdtek, akik a szóbeli tanúvallomást videóbizonyítékkal együtt hallják, 650%-kal nagyobb valószínűséggel őrzik meg az információt. Az audiovizualitás tehát erőteljesen befolyásolja az emberi emlékezetet és a valóság érzékelését, márpedig az internet tovább növelte a képek és a videók jelentőségét, mint tényszerű információk forrásait.[75]

Egyre több ember szerez híreket a közösségi médiából, így a nyomtatott sajtó lassan jelentőségét veszti. Ám a közösségi felületek teret adnak a deepfake-nek és terjesztik is azt, aminek következménye, hogy a közvélemény egyre szkeptikusabbá válik az audiovizuális tartalmak hitelességét illetően, vagyis a bíróság még a valódi hang-és képanyagban is kételkedhet, bizalma meginoghat még abban az esetben is, ha hitelesítik az adott bizonyítékot. Mindemellett, ha a hitelességet akár a vád, akár a védelem megkérdőjelezi, és stratégiáját erre építi, akkor a bizonyíték elveszítheti meggyőző erejét, amely által az igazságszolgáltatás intézményeibe vetett közbizalom is megrendülhet.[76]

6.2. Javaslatok a büntetőeljárás során felmerülő problémák megoldására

A deepfake igazságszolgáltatásra gyakorolt hatása 2018 közepén került a közvélemény figyelmének középpontjába. A megjelenő javaslatok és kommentárok hangsúlyozzák, hogy a mélyhamisított bizonyítékok jelentős kihívások elé állítják a bíróságokat, így sürgetik az érvek alátámasztásához használt audiovizuális anyagok hitelesítési procedúrájának minőségi ellenőrzését. Egyesek nehezményezik azt is, hogy a bírósági eljárások során a deepfake

[73] Delfino, Rebecca A.: i.m. – 310.
[74] Delfino, Rebecca A.: i.m. – 310.
[75] Delfino, Rebecca A.: i.m. – 311.
[76] Delfino, Rebecca A.: i.m. – 312.

bizonyítékok mérlegelésénél figyelmen kívül hagyják az olyan tényeket, mint a „hazugok osztaléka” vagy a bírói szkepticizmus. Az effajta kellemetlenségeket leginkább a bizonyítékok hitelesítésére vonatkozó szabályok módosításával, a bírók (egyes jogrendszerekben az esküdtek) továbbképzésével, valamint a mélyhamisítás védelmének korlátozásával lehetne kezelni. A probléma eljárási aspektusairól azonban még nem született átfogó, tudományos kutatás.[77]

A bizonyítás terén a legalapvetőbb nehézséget a digitális anyagok eredetiségének igazolása jelenti, mivel a vizsgálathoz szükséges szakértői elemzés kidolgozása időigényes és igen költséges lehet. Megoldásként szolgálhatna, hogy nem minden kérdésben értékelnék a digitális kép- és hanganyagokat, hanem többféle eszközre támaszkodnának a bizonyítékok hitelességének megállapítása során, ideértve a forrásokat, a szemtanúk általi megerősítést, valamint a többi bizonyítékok hitelességét, tehát mindent egybevetve mérlegelnének. A technológia fejlődése pedig hamarosan lehetővé teszi, hogy minden eszközön elérhetővé váljon az önhitelesítő szoftver, amely ugyanúgy segítséget nyújthat az eljárások során.

Ezt megelőzően már hangsúlyoztam az audiovizualitás szerepét. A bírói gyakorlatot vizsgáló amerikai empirikus felmérések alapján a bírák nem hajlamosak olyan mértékben a kognitív hibákra, mint a laikus esküdtek, bár tény, hogy ők sem immunisak teljes mértékben e hatásokra.[78] Fennáll ugyanis annak a veszélye, hogy az esküdtek a „seeing is believing” szemléletmód alapján előítéleteket kreálnak, amely befolyásolja a tanácskozási folyamatot.[79] A rendelkezésre álló, korlátozott számú kutatások alapján elmondható, hogy míg az esküdtek akkor sem képesek felismerni a mélyhamisításokat, ha azoknak közvetlenül ki vannak téve, addig a bírói tapasztalat bizonyos védelmet nyújt a kognitív torzítások ellen. Ennek oka, hogy a deepfake megjelenése előtt is csak a racionalitással, a szkepticizmus visszaszorításával, valamint a lényeges tények fókuszba helyezésével tudtak megfelelően mérlegelni. Következtetésképp megállapítható, hogy ahogy a kontinentális jogterületeken, úgy az angolszász jogban is növelni kell a bírók ténymegállapító funkcióját, továbbá mindkét jogrendszer esetében segítséget nyújthatna a bírák továbbképzése a digitális kriminalisztikai technológia terén, amely csökkenthetné az ügyekben felmerülő esetleges elfogultságokat és észlelési hibákat.[80]

[77] Delfino, Rebecca A.: i.m. – 340.
[78] Delfino, Rebecca A.: i.m. – 347.
[79] Delfino, Rebecca A.: i.m. – 346.
[80] Delfino, Rebecca A.: i.m. – 347.

7. A jogi szabályozás további lehetőségei és a felelősség megállapítása

Jelen állás szerint a technológia fejlődésével a jog nem lesz képes az MI-rendszerek minden aspektusát szabályozni, vagyis olyan sajátos szabályozási egységben kell gondolkodnunk, amely elkülönülten foglalkozik a problémával és lép fel ellene. Bár a hatalmas számítógépkapacitás és az összetett feladatok megoldása nem indokolná az önálló szabályozást, az MI általi kihívások nem ebből fakadnak: más technológia használatával együtt a kockázatok felerősödnek. Ami pedig a felerősödést okozza, az a bármely más műszaki eszköznél nagyságrendekkel nagyobb döntéshozatali önállóság, amely az MI-t jellemzi. Ezen autonóm és kognitív funkciók (pl. tapasztalatból történő tanulás) fejlődésével az MI-rendszerek eljuthatnak arra a pontra, amikor az eszköz olyan nagyfokú önállóságot mutat, hogy azt egy önállóan cselekvő tényezőnek kell tekintenünk, ezzel együtt pedig már a felelősségre vonás gyakorlati kérdései is jelentőssé válnak.[81] A jogalkotók hamar szembesülhetnek azzal a dilemmával, hogy eszközként vagy ágensként kezeljék az MI-t, és döntésük függvényében megalkossák a felelősség fokozatait. A logikát követve eszközalapú szabályozás esetén a felelősség a tervezőkre hárulna, míg, ha ágensként tekintünk rá, az utasításokat adó és működést ellenőrző felhasználóé vagy üzemben tartóé a felelősség. Ezzel együtt felmerül azon igény is, hogy a gyártók úgy alkossák meg MI alapú eszközeiket, hogy az átlagos felhasználók is különleges elővigyázatosság nélkül legyenek képesek biztonságosan és ellenőrizhető módon üzemeltetni azokat.[82]

Hosszú évek után, nagyjából egy időben született meg kettő fontos jelentés: az egyik az Amerikai Egyesült Államokban, a másik az Európai Unióban. Az USA elnöki hivatalának tudományos és technológiapolitikai irodája az MI jövőjére való felkészülésről szóló jelentését 2016 októberében adta ki, míg az EU-ban 2017 februárjában fogadták el a robotikára vonatkozó, polgári jogi szabályokról szóló Delvaux-jelentést. A két dokumentum közös pontja, hogy az MI társadalomban betöltött helyét és szerepét igyekeznek előre jelezni, azonban más-más problémákra keresnek megoldást. Így mindkét jelentés tartalmaz olyan javaslatokat, amelyekről elképzelhető, hogy formálni fogják az MI-ről kialakított gondolkodást.[83] Az amerikai javaslat az innovációt helyezi előtérbe, és kifejti azon álláspontját, miszerint az MI hozzájárulhat egyes társadalmi problémák megoldásához. A jogalkotó szabályozó szerepét

[81] G. Karácsony Gergely: *Okoseszközök – Okos jog? A mesterséges intelligencia szabályozási kérdései*, Ludovika Egyetemi Kiadó, Budapest, 2020, 69.
[82] G. Karácsony: i.m. – 70.
[83] G. Karácsony: i.m. – 71.

korlátozná, és kijelenti, hogy a hatalom birtokában a kormánynak elsődlegesen arra kell törekednie, hogy ne akadályozza beavatkozásokkal az innovatív fejlesztéseket. A jelentés továbbá külön felhívja a figyelmet az MI munkaerőpiacra gyakorolt hatására, mivel szerinte a folyamat egyaránt munkahelyeket fog teremteni és megszüntetni.[84] Az etikai kérdésekre speciális javaslatot nem tesz, de kiemelt témaként kezeli a nyilvános adatot tartalmazó bázisok szükségességét, az adatvédelmi kérdéseket, valamint az MI fejlesztésével foglalkozó mérnökök etikai képzését. Ez a megközelítés azonban azt a veszélyt rejti magában, hogy a kormány teljes egészében áthárítja az MI etikus működésének és tervezésének felelősségét a polgárokra és a magánszektorra.

Ezzel szemben az Európai Parlament által elfogadott jelentés inkább a robotikára koncentrál, az MI-t pedig nem önálló technológiaként, hanem a robotikai megoldásokban alkalmazott eszközként tartja számon. A kizárólag szoftver alapú, fizikai oldallal nem rendelkező MI-rendszert teljes egészében kihagyták a jogalkotók. A jelentés központi elemét képezik a robotika munkaerőpiacra gyakorolt hatásai, azok közül is annak vizsgálata, hogy a munkaerőpiaci változásokat előre jelző mechanizmusokat hogyan lehetne még inkább fejleszteni, illetve megfelelően alkalmazni.[85] Hosszú távú problémaként azonban felveti, hogy a munkakörök automatizálásával egyre kevesebb potenciális adófizető lesz, az állami bevételek tehát csökkeni fognak. Ennek orvoslására egy új adó kivetését sem veti el a jelentés, és emellett a robotizációra áttérő cégeket a társadalombiztosítási hozzájárulás összegének nyilvánosságra hozatalára kötelezné. Ezenkívül egy új szervezet létrehozását is szorgalmazzák a jogalkotók: a robotikával és MI-vel foglalkozó európai ügynökség feladata a piaci és technológiai újítások követése, a jó gyakorlatok felkutatása és a szabályozási lehetőségekre való javaslattétel, illetve a fogyasztóvédelemre vonatkozó kérdések kezelése lenne. Az ügynökség vezetné az intelligens robotok nyilvántartását is, amelybe regisztrálnia kellene minden olyan szervezetnek vagy személynek, aki ilyen roboteszközöket használ. Általános elvként megfogalmazza, hogy a robotok és MI-rendszerek fejlesztése és alkalmazása során az embereket nem lehet kitenni nagyobb kockázatnak, mint aminek a mindennapi életük során ki vannak téve.[86] A távoli jövőre nézve a jelentés elképzelhetőnek tartja a kifinomult autonóm robotok számára létrehozott elektronikus személyiség kialakítását, jogokat és kötelezettségeket megállapítva ezzel (többnyire az általuk okozott kárért való felelősség körében). Összegezve a jogi szabályozás és a soft law tárgykörébe tartozó eszközök együttes alkalmazását javasolja, előtérbe helyezve az

[84] G. Karácsony: i.m. – 71.
[85] G. Karácsony: i.m. – 72.
[86] G. Karácsony: i.m. – 73.

uniós szintű szabályozást.[87] A jogalkotók kiemelten fontosnak tartják még a roboteszközök valós körülmények között való tesztelését, hogy az esetleges hibák és problémák a felszínre kerüljenek.[88]

Felmerülhet azonban a kérdés, hogy a vajon a büntetőjogi és polgári jogi szabályozás, illetve a felelősség megállapítása elrettentően hatna-e az esetleges visszaélések tekintetében és orvosolná-e azokat a károkat, amelyeket a deepfake-ek okoznak? A teljes körű tilalom eleve kizárt, hiszen a mélyhamisítványok bizonyos esetekben súlyos károkat idézhetnek elő, de nem mindenben. A deepfake betiltása megakadályozná a digitális tartalmak egyértelműségét és minőségét javító rutinszerű módosításokat, ami pedig elvetné a művészetben, oktatásban és tudományban való fejlődés lehetőségét.[89]

Egyfajta megoldásnak tűnik magának az alkotónak a perlése, ez esetben azonban több jogterület is felmerülhet, többek között a szellemi tulajdonjog és a kártérítési jog. Ha szerzői jogból indulunk ki, akkor megállapíthatjuk, hogy egyes hamisítványok szerzői jogvédelem alatt álló tartalmakra épülnek, ami lehetőséget nyújt a pénzbeli kártérítésre, illetőleg a jogsértő tartalom eltávolításának követelésére. A bíróságnak azonban ebben az eseteben is meg kell állapítania, hogy a deepfake oktatási, művészeti vagy önkifejezési célból készült-e a szerzői joggal védett anyag felhasználásával. A „tisztességes felhasználás" pedig erősen tényspecifikus vizsgálat, amelyre vonatkozóan még nincs gyakorlati tapasztalat.[90] A másik lehetőség magának a tartalomszolgálatónak a perlése, hiszen kulcsszerepet játszanak a deepfake-ek terjedésében, illetve a hamisítványok alkotóit bizonyos esetekben nehéz megtalálni, éppen a közösségi oldalak szerteágazó volta miatt. Ezért a leghatékonyabb módnak a platformok felelősségre vonása lehet. Az online oldalakat már most is ösztönzik a tartalmak szűrésére, többek között erkölcsi alapú érvekkel, a piaci dinamika és a politikai nyomás hatásával.[91] Egy ésszerű szabályozás kellő gondossággal és biztosítékokkal tehát képes lenne a lehetséges visszaéléseket csökkenteni anélkül, hogy indokolatlanul akadályoznák az internet technikai fejlődését. Ezzel együtt elkerülhető lenne az is, hogy a platformok akaratlanul is deepfake tartalmak terjesztői legyenek, hivatkozási alapot szolgáltatva a károsultaknak. A megfelelő szabályozás egyúttal

[87] A soft law fogalmát tekintve a gyakorlatban két meghatározás alakult ki. Az egyik szerint a kötelező egyezményeken belüli bizonytalan tartalmú szabályok tekinthetők nemzetközi puha jognak. Egy másik megközelítés értelmében a soft law-t olyan nemzetközi szabályok alkotják, amelyek az alkotók szándéka szerint nem hoznak létre nemzetközi jogi kötelezettséget.
[88] G. Karácsony: i.m. – 74.
[89] Citron – Chesney: i.m. – 1792.
[90] Citron – Chesney: i.m. – 1793.
[91] Citron – Chesney: i.m. – 1795.

lehetővé tenné, hogy könnyebben különbséget tegyünk az ártalmatlan és a jogellenességet gerjesztő weboldalak között. Be kell látnunk azonban, hogy csak és kizárólag ez a lépés kevés lenne, hiszen a platformokon jelenlévő obszcenitás, zaklatás vagy erőszak más szakpolitikai válaszokra is ösztönöz.[92]A jogalkotónak ebből kifolyólag egy olyan szabályozási optimumot kell kifejlesztenie, amely biztosítja az emberek biztonságát és jogait, viszont nem akadályozza az új technológiák nyújtotta előnyök kereskedelmi felhasználását, azok fogyasztói élvezetét és lehetőleg előmozdítja az innovációt is.[93]

Feltétlen figyelmet kell szentelni a szabályozást megalapozó adatokra, azok körére és forrására. Egyrészt akadályt jelenthet az elégtelen mennyiségű adat vagy minta megléte, másrészt pedig az adatok helyességét és az azokból elvont következtetéseket bíráló szakemberek közötti konszenzus hiánya. Problémát jelenthet még az is, hogy a releváns adatokat az új technológiákban valamilyen módon értintett érdekcsoportok igyekeznek elfedni, megmásítani, befolyásolni. Végsősoron számolni kell azzal, hogy a releváns információk egy része valószínűsíthetően egyelőre sem a fejlesztők, sem a felhasználók, sem a jogalkotók előtt nem ismert. Ilyenek lehetnek az újdonsült technológia jövőbeni felhasználási lehetőségei vagy a használatból fakadó későbbi következmények.[94] Az elmúlt évtizedekben a technológiai fejlődés olyan szinten felgyorsult, hogy a jogalkotó nem tudja követni az évtizedekkel ezelőtt még adekvátnak számító jogalkotási módszertant. Így könnyen abba a dilemmába eshet, hogy döntenie kell: meggondolatlanul és gyorsan cselekszik, vagy inkább a teljes bénultságot választja és bezár a probléma előtt. Ilyenkor általában az óvatos megközelítés győz, a jogalkotó elővigyázatosságból egy status quo-t fenntartó, korlátozó jellegű intézkedést hoz, ami azonban megakadályozza az új fejlesztések hatékony és időben való piacra jutását. Globális szempontból ez jelentős hátrányt jelenthet, hiszen alulmaradáshoz vezethet a fejlesztéseket támogató országokkal szemben.[95]

A felmerülő károkkal kapcsolatban a jogalkotó négy alapvető felelősségi helyzetet állapíthat meg. Az első esetkörbe tartozik a gyártó hibájából vagy az általa elkövetett hiányosságból fakadó károkért való felelősség. Ez abban az esetben lehetséges, ha az MI-t alkalmazó termék gyártási hibájából egyenes következik a kár vagy a jogsértő eredmény.[96] Előfordulhat, hogy az üzemben tartó szakszerűtlen vagy szabálytalan magatartásából vagy az

[92] Citron – Chesney: i.m. – 1800.
[93] G. Karácsony: i.m. – 75.
[94] G. Karácsony: i.m. – 76.
[95] G. Karácsony: i.m. – 77.
[96] G. Karácsony: i.m. – 105.

általa ellátott felügyeleti tevékenység nem megfelelő elvégzéséből fakad a kár vagy jogsérelem. Ez esetben a veszélyes üzemi felelősség szabályai alkalmazandók, feltéve, ha a kár az MI kiszámítható működéséből ered, tehát annak bekövetkezte racionálisan valószínűsíthető volt. Elképzelhető továbbá szándékos károkozás is, amikor a gyártó vagy az üzemben tartó személy, esetleg egy rosszakaró olyan módon módosítja az MI működését, hogy az ésszerű keretek között előrelátható módon kárt vagy jogsértést eredményez, és a sérelem a módosítás nélkül nem következett volna be. Adott esetben a felelősség közvetlenül a kár bekövetkezését okozó, módosítást végrehajtó személyt terheli mind polgári jogi, mind büntetőjogi tekintetben.

A fentebb sorra vett esetek a felelősségi alakzat szempontjából feltételezik, hogy az MI-n alapuló algoritmus működése kiszámítható, átlátható és előre meghatározható. Ám vannak esetek, amikor az MI autonóm módon hozott döntésére senki nincs közvetlen hatással, mivel az nem következik egyenest a programból, így bekövetkeztének lehetősége sem állapítható meg, időpontja, módja nem előre látható vagy kiszámítható, illetőleg a döntési mechanizmus a felügyeletet gyakorló számára nem átlátható. A büntetőjogi felelősség ez esetben nem állapítható meg, hiszen hiányzik a szándék és a cselekmény fogalmi eleme a tényállásból; ennek ellenére nem tekinthetünk el a ténytől, hogy bűncselekmény valósult meg. A büntetőeljárásban ezt alapul véve kell megakadályozni, hogy az MI-rendszer újabb bűncselekményt tudjon elkövetni, vagyis a büntetőeljárásban olyan megelőző adminisztratív és műszaki intézkedések szükségesek, amelyek jelenthetik a konkrét MI működésének megszüntetését, újra programozását, a gyári beállítások visszaállítását vagy a törlés utáni ellenőrzött újra tanítást.[97]

A nehézség megoldása azonban nem csak a jogi szabályozás mentén kereshető. A médiaoktatás és a kritikus gondolkodás elősegítése pozitív hatással lehet a jogi szabályozás módszertanára is[98], mivel gyorsíthatja annak társadalom általi adaptálását. Kiemelendő emellett a fact-checkerek („tényellenőrök") szerepe is, amelyek rövid időn belül ellenőrizni tudnák a deepfake tartalmak eredetiségét.[99] Ezenfelül a magánszektor is reagálhat a problémára: olyan szolgáltatásokat hozhatnak létre, amelyek célja, hogy megvédjék az ügyfeleket a hamisításon alapuló károk bizonyos formáitól. Ilyen fejlesztések leginkább a személyazonosság-lopás által okozott vesztségekre válaszul jelenhetnek meg.

[97] G. Karácsony: i.m. – 106.
[98] Gosztonyi – Lendvai: i.m. – 46.
[99] Gosztonyi – Lendvai: i.m. – 46.

8. Összegzés

A deepfake jelenlétével számolni kell, eltüntetni teljes mértékben azt nem lehet, legfeljebb korlátozni a hozzáférhetőséget és megszabni azokat a célokat, amelyekre kifejezetten engedélyezett a felhasználás. Véleményem szerint az egész emberiség rendkívül nagy kihívás előtt áll, amire nem lesz egyszerű konszenzusos megoldást találni, egyöntetűen reagálni. Az ideálok közvetítésével, gyakorlatilag a való élet határain való átlépéssel az MI egyre csak átformálja az emberek mindennapi gondolkodását, ezáltal az egész társadalmat. Minél előbb cselekedni kell, hiszen a hamisítások megállíthatatlanul terjednek, ráadásul nem túl ígéretes a jövőre nézve az sem, hogy a deepfake létrehozására alkalmas applikációk és programok bárkinek a rendelkezésére állnak: nem köthetők adott elkövetői csoportokhoz, mintázatokhoz, helyszínekhez, időpontokhoz. A deepfake a világon mindenütt és folyamatosan jelen van, terjedését pedig elősegíti azon előnye, hogy a hagyományos vizuális manipulációval szemben sokkal kevésbé időigényes és jóval kevesebb hozzáértést, valamint befektetett energiát követel. A digitális manipulált tartalmak olyan idealizált szituációkat lebegtetnek az emberek előtt, amelyek teljesen irreális elvárásokat támasztanak. Ezáltal a valóságot egyre kevésbé fogadják el, befelé fordulnak, társaikkal megszűnik a mindennapi közvetlen kommunikáció. Ezt a folyamatot egyáltalán nem segítik elő az interneten burjánzó álhírek és a propaganda, amelyek megosztottságot eredményeznek, a közbizalmat és a köznyugalmat megrendítik. A demokratikus rendszer, ezzel az igazságszolgáltatás is, mint hatalmi ágak egyike, veszélybe kerül. Ez kihatással van az állampolgárokra, a politikai folyamatokra, a nemzetközi kapcsolatokra, tehát mindenre, amely egy állam működésének alapját képezik.

A deepfake-ek büntetőeljárásban való kezelésének központi eleme a bizonyítékok hitelességének biztosítása. Feltétlenül szükséges a digitális audiovizuális anyagok ellenőrzésére szolgáló hitelesítési eljárások fejlesztése, mivel a jelenlegi technológiák nem mindig képesek különbséget tenni a valódi és hamisított tartalmak között. Ennek érdekében szigorítani kell a büntetőeljárásban a bizonyítékok hitelesítési szabályait, hogy azok növeljék az igazságszolgáltatásba vetett bizalmat, és megakadályozzák, hogy a deepfake-ek megtévesszék a bíróságot. Emellett szükséges a nyomozók, az ügyészek és a bírók továbbképzése is, hogy jobban megértsék és felismerjék bizonyítékok deepfake jellegét, valamint fejlesszék a büntetőeljárásban felmerülő manipulált tartalmak kezeléséhez szükséges ismereteiket. Ezenfelül nagyon fontos a deepfake technológia jogi szabályozásának modernizálása, különös tekintettel az EU szabályaihoz való harmonizálása, az emberközpontú és megbízható MI elterjedését elősegítése érdekében. Az ilyen szabályozás célja, hogy csökkentse a technológia

káros hatásait, ugyanakkor egyértelmű kereteket biztosítson annak felhasználására, megvédve ezzel a társadalom tagjainak alapvető jogait és az igazságszolgáltatás integritását.

9. Felhasznált irodalom

9.1. Szakirodalom

Aczél – Veszelszki: Deepfake: A valótlan valóság, Gondolat Kiadó, Budapest, 2023.

Citron, Danielle K. – Chesney, Robert: Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, Berkeley School of Law, University of California, 2019.

Davis, Rama – Wiggins, Chris – Donovan, Joan: Deepfake, Belfer Center, Cambridge, 2020.

Delfino, Rebecca A.: Deepfakes on Trial: A Call To Expand the Trial Judge's Gatekeeping Role To Protect Legal Proceedings from Technological Fakery, Hastings Law Journal, San Fransisco, 2023.

Feeney, Matthew: Deepfake Laws Risk Creating More Problems Than They Solve, Regulatory Transparency Project of the Federalist Society, 2021.

Frenda, Steven J. – Knowles, Eric D. – Saletan, William – Loftus, Elizabeth F.: False memories of fabricated political events, Legal Studies Research Paper Series, University of California, 2012.

G. Karácsony Gergely: Okoseszközök – Okos jog? A mesterséges intelligencia szabályozási kérdései, Ludovika Egyetemi Kiadó, Budapest, 2020.

Gerstner, Erik: Face/Off: „DeepFake" Face Swaps and Privacy Laws, Defence Counsel Journal, 2020/1.

Goggin, Benjamin: From porn to 'Game of Thrones': How deepfakes and realistic-looking fake videos hit it big, Business Insider, 2019.

Gosztonyi Gergely – Lendvai Gergely: Deepfake és dezinformáció, Médiakutató, Budapest, 2024.

Herke Csongor: Deepfake: áldás vagy átok? Jogi szabályozási szempontok, Pro futuro, Debrecen, 2023.

Köbis, Nils C. – Doležalova, Barbora – Soraperra, Ivan: Fooled twice: People cannot detect deepfakes but think they can, iScience, 2021.

Langa, Jack: Deepfakes, real consequences: crafting legislation to combat threats posed by deepfakes, Boston University Law Review, 2021/101.

Meskys, Edvinas - Liaudanskas, Aidas - Kalpokiene, Julija - Jurcys, Paulius: Regulating deep fakes: legal and ethical considerations, Journal of Intellectual Property Law & Practice, 2020/15.

Mráz, Attila: Deepfake, demokrácia, kampány, szólásszabadság, In: Török Bernát – Ződi Zsolt (szerk.): A mesterséges intelligencia szabályozási kihívásai, Ludovika Egyetemi Kiadó, Budapest, 2021.

Palmiotto, Francesca: Detecting Deep Fake Evidence with Artificial Intelligence. A Critical Look from a Criminal Law Perspective, SSRN, 2023.

Sloot, Bart van der – Wagensveld, Yvette – Koops, Bert-Jaap: Deepfakes: The legal challenges of a synthetic society, Tilburg University, Tilburg, 2021.

Sorbán Kinga: A bosszúpornó és deepfake pornográfia büntetőjogi fenyegetettségének szükségességéről, Belügyi Szemle, 2020/10.

9.2. Internetes források

https://bigthink.com/the-present/reality-vs-deepfake/

https://copy21.com/2020/03/hiba-a-matrixban-nicolas-cage-neo-szerepeben-avagy-a-deepfake-ai-videok-szerzoi-jogi-szemmel-1-resz/

https://digiday.com/media/the-wall-street-journal-has-21-people-detecting-deepfakes/

https://en.wikipedia.org/wiki/Center_for_Medical_Progress

https://en.wikipedia.org/wiki/Planned_Parenthood

https://hoaxes.org/photo_database/image/the_tydings_affair

https://index.hu/techtud/2023/06/18/mesterseges-intelligencia-d-id-csorba-david-hang/

https://medium.com/@divakar1591/autoencoders-6fab1a9a9f9c

https://onezero.medium.com/deepfake-music-is-so-good-it-might-be-illegal-c11f9618d1f9

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4384122

https://petapixel.com/2012/11/15/the-first-hoax-photograph-ever-shot/

https://riskhubgroup.co.uk/wp-content/uploads/2024/01/5-Ways-Cybercriminals-are-Weaponsing-AI-1.pdf

https://www.aspi.org.au/report/weaponised-deep-fakes

https://www.aspi.org.au/report/weponised-deep-fakes

https://www.bbc.com/news/technology-49570418

https://www.belfercenter.org/sites/default/files/2020-10/tappfactsheets/Deepfakes.pdf

https://www.britannica.com/technology/deepfake

https://www.businessinsider.com/deepfakes-explained-the-rise-of-fake-realistic-videos-online-2019-6

https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/

https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

https://www.dw.com/en/frequency-illusion-when-seeing-is-believing-not-fact/a-64838404

https://www.france24.com/en/live-news/20240729-musk-faces-criticism-over-deepfake-kamala-harris-video

https://www.huffpost.com/archive/ca/entry/lyrebird-helps-als-ice-bucket-challenge-co-founder-pat-quinn-get-his-voice-back_ca_5cd54688e4b07bc729767cdb

https://www.newsweek.com/when-photographs-lie-206894

https://www.npr.org/sections/thetwo-way/2018/01/30/581853255/hawaii-missile-drill-stated-this-is-not-a-drill-resulting-in-false-alert

https://www.nytimes.com/2015/08/28/us/abortion-planned-parenthood-videos.html

https://www.nytimes.com/2020/07/01/movies/deepfakes-documentary-welcome-to-chechnya.html

https://www.reddit.com/

https://www.si.com/nfl/2016/04/29/laremy-tunsil-draft-controversy-ole-miss-dolphins-twitter-instagram-hacked

https://www.theatlantic.com/news/archive/2017/06/qatar-russian-hacker-fake-news/529359/

https://www.theguardian.com/media/2004/feb/17/newmedia.pressandpublishing

https://www.theverge.com/2020/4/28/21240488/jay-z-deepfakes-roc-nation-youtube-removed-ai-copyright-impersonation

https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn

https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

https://www.youtube.com/watch?v=QiiSAvKJIHo

9.3 Felhasznált képek

https://hoaxes.org/photo_database/image/the_tydings_affair#google_vignette

https://www.researchgate.net/figure/An-image-splicing-example-32-a-the-spliced-image-showing-John-Kerry-and-Jane-Fonda_fig1_319534758