Médiakutató, 2024. tavasz
XXV. évf. 1. szám, 41–49. o.
10.55395/MK.2024.1.3

Deepfake és dezinformáció

Mit tehet a jog a mélyhamisítással készített álhírek ellen?*

Gosztonyi Gergely
Eötvös Loránd Tudományegyetem
Lendvai Gergely
Pázmány Péter Katolikus Egyetem

E tanulmányban a deepfake és az álhírterjesztés kapcsolatát tárjuk fel a jog eszközeivel. Előbb a fogalmi alapokat ismertetjük, majd a dezinformáció és a deepfake kapcsolatát mutatjuk be, végül áttekintjük a deepfake-kel készített álhírtartalmak kontextusában releváns szabályozásokat – különösen az amerikai és az európai szabályozási irányokat. Külön figyelmet fordítunk a dezinformáció jogi megítélésére a deepfake-technológia kontextusában, kiemelve a káros társadalmi és jogi folyamatokat. E tanulmány alapját a nemzeti és a nemzetközi szabályozások és a releváns szakirodalom feldolgozása, illetve a deepfake-dezinformáció okozta polémiákra adott megoldási javaslataink jelentik.

Kulcsszavak: AI Act, deepfake, dezinformáció, mesterségesintelligencia-szabályozás, szólásszabadság

Deepfake and disinformation

What can the law do about fake news created by deepfaking?

This paper explores the relationship between deepfake and fake news through the lens of the law. It is divided into three main conceptual parts; first, it looks into the conceptual basis, then it explores the relationship between disinformation and deepfake, and finally it offers an overview of the relevant regulation in the context of deepfake fake news content (in particular, the US and the European regulatory trends). The paper pays particular attention to the legal perception of disinformation in the context of deepfake technology, highlighting the harmful social and legal outcomes involved. It is based on an overview of national and international regulations and of the relevant literature and includes the authors' proposed solutions to the controversies generated by deepfake disinformation.

Key words: AI Act, AI regulation law, deepfake, disinformation, freedom of expression

* Ez az elemzés a Bolyai János Kutatási Ösztöndíj és a Kulturális és Innovációs Minisztérium ÚNKP-23-5, illetve ÚNKP-23-3 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

1. Bevezetés

Orbán Viktor, Magyarország miniszterelnöke közismerten nem kedveli az Európa felé irányuló migrációt, és szeretné megtartani a kontinens „keresztény jellegét" (Rankin 2018). Mindezek ellenére lehetne-e találni olyan videót, ahol azt mondja, hogy ő nagy rajongója a migrációnak, és legszívesebben lerombolná az összes határt és kerítést Magyarország körül, hogy mindenkit befogadhasson, aki segítséget kér, mert súlyos válságok miatt kellett elhagynia szülőföldjét? Ez most lehetetlennek tűnik. De tényleg lehetetlen lenne? Hiszen mindössze egy ingyenes programra, minél több fotóra, egy komoly számítógépre és kellő mennyiségű időre lenne szükség ahhoz, hogy mindez megvalósuljon. A kérdés így az, hogy megváltoztathatja-e a deepfake az életünket? Megváltoztatja-e azt, amit a saját szemünkkel látunk és elhiszünk? E tanulmányban arra a kérdésre keresünk választ, hogy miként tudja a jog felvenni a harcot a mesterséges intelligencia által generált valóság és a dezinformáció ellen.

Ahogy a sajtó és a média fejlődött az elmúlt években, és ahogy a közösségi média mindennapi életünk szerves részévé vált, egyre inkább elveszítjük a szerkesztett valóságba vetett bizalmunkat.[1] Az emberek egyre kevésbé bíznak a sajtóban, hiszen a cikkeket pillanatok alatt át lehet szerkeszteni online, az álhírek valósága velünk van – még ha a politikusok többsége a kifejezést (helytelen és káros módon) a „valami, amivel nem értek egyet" formában is használja. Számos kutatás sorolta fel azokat az eseményeket, amelyek erodálták a médiaiparba, sőt az újságírásba vetett közbizalmat. Farnaz Fassihi, a Wall Street Journal munkatársa szerint „a közösségi média térnyerése megnehezítette az átlagember számára, hogy különbséget tegyen az ellenőrzött tények és a félretájékoztatás között" (in Dickinson 2018).

Mindez ráadásul nem csak a politikában van így. Láttunk néhány felkavaró videót ismert színésznőkről, akik elég nyilvánvaló szexuális helyzetekben voltak, amelyekkel kapcsolatban szinte biztosak lehettünk benne, hogy nem történtek meg. Kiderült, hogy olyasmivel állunk szemben, ami szó szerint örökre megváltoztathatja az életünket: a mesterséges intelligencia eszközeivel könnyen lehet hamis pornófelvételt készíteni bárkiről,[2] ugyanakkor „a büntethetőségről, illetve a büntetendőségről szóló élénk szabályozási-jogpolitikai diskurzus" (Sorbán 2020) miatt az nem egyértelmű még, hogy a büntetőjogi eszközök külön bűncselekményként azonosítják-e ezt. Sorbán Kinga (2020) a bosszúpornó és a deepfake pornográf felvételek esetében szóba kerülő büntetőjogi tényállások kapcsán kiemeli, hogy a büntetőjog akár „a becsület csorbítására alkalmas hamis hang- vagy képfelvétel készítése, rágalmazás, szexuális kényszerítés, esetleg zsarolás" tényállásokat is alkalmazhatna, ugyanakkor „Magyarországon sem a bosszúpornó, sem a deepfake pornográfia nem jelenik meg önálló tényállásként a büntető törvénykönyvben".

A szexuális tartalom csak az első lépés. A következő a politika, majd az átlagemberek világa lehet. Ha elveszítjük a bizalmunkat az írott cikkekben, vajon a képekben és a videókban is elveszítjük a bizalmunkat (Szűts 2018)? Ha már nem hiszünk majd a saját szemünknek, a világ ismét nagyon más lesz.

2. Fogalmi keretek: mi a deepfake? Arcplagizálás, mélykamu és mélyhamisítás

A deepfake[3] mint elnevezés a *deep learning* (mély tanulás) és a *fake* (hamis) szó kombinációjának eredménye (Mráz 2021); a fogalom első eleme tehát a mély tanulásra, míg a második elem a gyártott tartalom hamisított jellegére utal (Citron & Chesney 2019). Bár a deepfake fogalmát számtalan szemszögből meg lehet közelíteni, a jog szempontjából érdemes Lendvai Gergely Ferenc (2023a) több forrásból szintetizált deepfake-definícióját felhasználni a konceptuális alapok lehelyezéséhez. A fogalom szerint a deepfake mesterséges intelligencián alapuló technológia, amely segítségével a felhasználó célzatosan és szándékosan hamisított, leginkább emberekről készített kép- és hangfelvétel-tartalmakat tud készíteni, oly módon, hogy a létrehozott tartalom meggyőző módon képes a hamisított tartalmat valóságosnak

1 Húsz ország átlagában tízből csak hat internetező (63%) mondta azt, hogy bízik az internetben. Ez 11 százalékpontos csökkenés a 2019-ben végzett hasonló felmérés óta (Ipsos 2022).

2 Például a Wonder Woman sztárjának, Gal Gadot-nak az arcát a mesterséges intelligencia segítségével egy pornósztáréra montírozták, így ő egy teljes filmben „szerepel". Mellette Scarlett Johansson, Maisie Williams, Taylor Swift vagy Aubrey Plaza is áldozatául esett már hasonlónak (Cole 2017).

3 Az angol deepfake kifejezés a magyar nyelvben lehet főnév és melléknév is.

érzékeltetni. A hamisított tartalmak tehát – ahogy Aczél Petra (2023) is kiemeli – leegyszerűsítve képmásvisszaélések, arcplagizálások, amelyek az emberi tulajdonságokba vetett bizalmat kérdőjelezik meg.

A deepfake története akkor kezdődött, amikor 2016-ban arccserélő alkalmazások kezdtek elterjedni a közösségi médiában. Bár a technológia 2013-ból származik, három évbe telt, hogy az olyan alkalmazások, mint az MSQRD, a Face Swap Live, a Snapchat és a Face Stealer a szélesebb közönség számára is elérhetővé váljanak (Dredge 2016). Egy ideig sokak közösségi médiában megjelenő üzenőfala tele volt ilyen képekkel. A fejed egy baba testén, egy macska testén vagy a szerelmed testén. Néha vicces, de többnyire undorító eredményekkel. Nem volt más, mint gyorsan múló érdekesség. És amilyen gyorsan jött, olyan gyorsan el is tűnt.

2017 végén azonban egy sokkal riasztóbb dolog történt: a Motherboardon egy DeepFakeApp nevű felhasználó lehetővé tette, hogy valaki fejét rámontírozza egy mozgóképre, azaz egy teljes videóban szerepeltesse (Vincent 2017). A technológia és az emberi kreativitás innen villámgyorsasággal kezdett fejlődni. A fentieket példázva klasszikus deepfake tartalom lehet egy videó, amelyen valaki fejét egy másik személy testére montírozzák, például Jim Carrey fejét Jack Nicholson feje helyére a *Ragyogás* című film ikonikus jelenetében (Spellberg 2019).

A technológiai részletek terén számtalan deepfake-készítési lehetőség van, a legismertebb a „GAN", azaz Generative Adversial Networks technológia (Pantserev 2020). Ez a technológia egy olyan neurális hálózatot, generátort takar, amely adathalmazok és adattartalmak letöltésével készít új, hamisított tartalmakat. Ez annyit tesz, hogy bizonyos kép- és hanganyag betáplálásával (Reid 2021: 209) a deepfake-technológia képes rámontírozni az általunk szolgáltatott anyagokat egy másik, különböző tartalomra, tehát

> …egy személyről rendelkezésre álló képek vagy hangfelvételek alapján olyan manipulált felvételt készít, amely a valós személyt egy olyan fiktív helyzetben ábrázolja, amilyen helyzetben a konkrét személy nem szerepelt, illetve olyan kijelentést tulajdonít neki, amilyen kijelentést ő nem tett meg (G. Karácsony 2022: 300).

Mitől lesz tehát egy deepfake tartalom dezinformációt érintő kérdés? Érvelésünk sarokköve az, hogy a deepfake szólás, és mint ilyen, ugyanúgy érdemes és kell is a deepfake szólásszabadsági (Lendvai 2023a) és álhírterjesztéssel kapcsolatos polémiáiról is beszélni. E polémiák pedig magukkal vonzzák a jog lehetséges válaszait és a szabályozási megoldások, illetve legjobb gyakorlatok kérdéseit.

3. Deepfake és dezinformáció

3.1. Átlátszó hack vagy dezinformációs nehézfegyver?

Volodimir Zelenszkij ukrán elnök 2022 márciusában beszédet mond egy videón, amely futótűzként terjed az ukrán és orosz közösségi médiában – Zelenszkij a zavaros felvételen bizarr arcmozgással felszólítja az ukrán katonákat, hogy tegyék le a fegyvert (Allyn 2022). Joe Biden amerikai elnökről 2023 májusában egy olyan videó kerül fel a Facebookra, amelyen az elnök az unokája mellkasát fogdossa (Lima 2023). Tartalmak, amelyek elérhetőek, láthatóak, valóságosnak tűnőek – mégis hamisak.

Rob Cover (2022: 615) szerint a deepfake elsősorban „társadalmi aggály". E társadalmi aggály pedig elsősorban Adam Satarino és Paul Mozur cikkének (2023) tételmondatában áll: az emberek hamisak, de a dezinformáció valós. A deepfake társadalmi aggályként való értelmezését könnyen alá lehet támasztani empirikus kutatási eredményekkel is, hiszen a politikai mikrotargetálásnál hatásosnak bizonyulhatnak a deepfake tartalmak (Dobber et al. 2020). A deepfake tartalmak ezen felül mind a jog, mind a demokratikus nyilvánosságot érintő kérdések területén számtalan ellentmondásos hatást váltanak ki; személyiségi és arcképfelhasználási polémiákat okoznak (izgalmas gondolati ívet vázol a deepfake és a karaktergyilkosság kapcsolatáról Bajomi-Lázár Péter [2022: 8]), és képesek negatívan befolyásolni a közvélemény alakulását (Van der Sloot & Wagensveld 2022). E körben kiemelendő a fentebbi bekezdésben említett politikai megtévesztésre vagy akár a fegyveres konfliktusokban történő deepfake-alkalmazás is (Allen 2022). Andrew Ray aláhúzza (2021), hogy a deepfake tartalmak nagyban hozzájárulhatnak ahhoz, hogy csökkenjen a politikai bizalom, és valós veszélyt hordozhatnak a választásokba való illegális beavatkozások terén is.

Veszelszki Ágnes (2022: 33) szavaival élve a deepfake tartalmak tekintetében „igazi nehézfegyverzetről" beszélünk, hiszen a hamis hírek elleni harc egyik legnagyobb kihívása az álhírek készítésén túl az is, hogy a deepfake által generált fake news-t a legtöbbször nem is a dezinformációs szándékkal terjesztő felhasználók terjesztik a legnagyobb mértékben, hanem azok, akiket megtévesztett a manipulált tartalom. E tekintetben megkerülhetetlen Crisrian Vaccari és Andrew Chadwick (2020) kutatásának megemlítése, amely az elsők között mérte fel, hogy mennyire megtévesztőek az akkor elérhető deepfaketechnológia-tartalmak. A kutatásból kiderül, hogy egy 4 másodperces deepfake videó a kutatásban résztvevők közel 15 százalékát tévesztette meg, míg 35 százalékuk bizonytalan volt azzal kapcsolatban, hogy a tartalom, amelyet láttak, „valós" volt-e. Ez az arány romló eredményt mutatott egy 26 másodperc hosszúságú deepfake videónál, ahol a megtévesztettek aránya 16 százalék, míg a bizonytalanok aránya majdnem 37 százalék volt. Itt kell kiemelni, hogy a deepfake tartalmak „valóságossága" kapcsán nem feltétlenül a megtévesztettek aránya a vészjósló, jóval inkább a bizonytalanoké – ha a három évvel ezelőtti deepfake-technológia ugyanis az emberek több mint felében bizonytalanságot, rosszabb esetben „átverést" generált, joggal tehetjük fel, hogy a 2023-ban elérhető deepfake-technológia jóval nagyobb arányban lenne képes megtéveszteni a tartalomfogyasztót, és bizonytalanságot okozni benne.

3.2. Deepfakes Act és AI Act – a deepfake tartalmakra és álhírekre adott jogalkotói válaszok

Andrew Ray (2021: 991–992) három szabályozási problémát vázol fel a megtévesztő politikai deepfake-ek szabályozása kapcsán. A szerző szerint az első probléma az, hogy a deepfake tartalom eltávolítása nem jár együtt a deepfake tartalom káros hatásainak megoldásával – az eltávolítás ugyanis nem eredményezi azt, hogy a káros hatás elszenvedőjének jogi helyzete javulna vagy reputációja helyreállna. A második probléma az, hogy nem mindegyik közösségimédia-platform kezeli a dezinformációs tartalmak körében a deepfake-et. A harmadik probléma pedig a deepfake konceptualizációjára vonatkozik: a deepfake pontatlan fogalma miatt a jog nem tudja hatékonyan megvédeni az érintetteket. E három problémakör mentén az alábbiakban bemutatjuk, milyen szabályozási irányok figyelhetőek meg – elsősorban az Egyesült Államokban és az Európai Unióban.

3.2.1. Amerikai megoldások

Az amerikai deepfake-szabályozások kapcsán fontos röviden értekezni a különböző és gyakran eltérő tagállami gyakorlatokról. Jelenleg nincs az Egyesült Államokban szövetségi szintű szabályozás a deepfake tartalmak káros kezelése kapcsán, ugyanakkor számos tagállam rendelkezik egyedi, egyes esetekben pedig speciális tárgyú (csak a választásokkal kapcsolatos) deepfake szabályozásokkal.

Elsőként érdemes egy kezdetleges szövetségi próbálkozást megemlíteni. 2019-ben Yvette Clarke demokrata képviselő benyújtotta a Kongresszusnak a DEEPFAKES Accountability Act névre hallgató első, tagállami szabályozások feletti deepfaketörvény-tervezetet (Kocsis 2022, H.R. 3230 § 1041). A törvény lehetőséget nyújtott volna arra, hogy kártérítést követelhessenek a káros deepfake tartalommal érintett felek, illetve egységes deepfake definíciót is proponált (H.R. 3230 § 1041(n)(3)):

> A deepfake olyan videofelvétel, mozgóképes film, hangfelvétel, elektronikus kép vagy fénykép, illetve a beszéd vagy magatartás bármely olyan technológiai ábrázolása, amely A) úgy tűnik, hogy hitelesen ábrázolja egy olyan személy beszédét vagy magatartását, aki valójában nem vett részt ilyen beszédben vagy magatartásban; és B) amelynek előállítása valójában technikai eszközökön múlik, és nem pedig egy másik személy azon képességén, hogy fizikailag vagy szóban megszemélyesítse az ilyen személyt.

A feltételes múlt idő használata ugyanakkor nem véletlen: Clarke javaslata megbukott kongresszusi szinten. Az amerikai szövetségi szintű deepfake-szabályozási igény ugyanakkor nem szűnt meg: Joe Biden e tanulmány

megírása előtt pár héttel, 2023 októberében aláírta a mesterséges intelligenciáról szóló végrehajtási rendeletet (Executive Order), amely a világon elsőként kíván megoldást nyújtani a mesterséges intelligencia, így a deepfake okozta problémákra (Johnson 2023). Ez utóbbi kapcsán érdemes felidézni Biden reakcióját is, amikor megnézett egy róla készült deepfake videót. 2023. október végén egy sajtótájékoztatón ugyanis az amerikai elnök kifejezte aggodalmát a káros AI-tartalmak kapcsán, és megjegyezte, hogy egy róla készült deepfake videó őt is megtévesztette, és az első reakciója az volt, hogy „mikor mondtam én ezt?" (Yahoo UK 2023).

Az amerikai tagállami szabályozások közül kiemelkedik Kalifornia esete, ahol két speciális deepfake-törvényt lépett életbe 2020-ban: az egyik a kampányanyagok deepfake-technológiával való hamisítására, míg a másik szabályozás magánjogi kereseti jogot biztosít olyan mesterséges intelligenciával készített tartalom előállítójával szemben, aki szándékosan és tudatosan szexuális tartalmat hoz létre másról, annak beleegyezése nélkül. A kampányanyagok kapcsán a kaliforniai választási törvénybe iktatták be a 20010. szakaszt, amelynek értelmében egy politikai tisztséget betölteni kívánó jelölt a választást követő 60 napon felléphet azzal a deepfake tartalmat gyártó személlyel szemben, aki kvázi „le akarta járatni" a jelöltet a kampány során a tartalommal (Van der Sloot & Wagensveld 2022: 11; Californian Elections Code: §20010). Rob Cover (2022) gondolataihoz csatlakozva e körben megállapítható, hogy a kaliforniai törvénymódosítás egy kritikus kérdés orvoslására irányul – ahogyan azt a 3.1. alfejezetben is több ízben bemutattuk, a rosszindulatú deepfake tartalmak képesek lehetnek nagymértékű kárt okozni egy politikai aspirációval rendelkező személy karrierjében. Kérdéses ugyanakkor, hogy a szabályozás – feltételezve, hogy végrehajtható, azaz megtalálják a lejáratót anyagot készítő deepfake tartalom előállítóját – orvosolja-e a lejáratókampánnyal érintett jelölt kárát.

Visszatérve Andrew Ray első problémájára: a kaliforniai szabályozás részmegoldást nyújt az érintettnek – ugyan lehetősége van a jelöltnek fellépni a tartalomelőállítóval szemben, a renoméját ért kár, illetve akár egy esetleges választási vereség esetén a rendelkezés már nem nyújt valós segítséget neki.

Távolabbi és általánosabb a kapcsolódás a dezinformáció-szabályozás és a deepfake között Virginia államban, ahol 2019-ben lépett hatályba a más személyről készült képek jogellenes terjesztéséről vagy értékesítéséről szóló törvény (Unlawful Dissemination or Sale of Images of Another Person) (Ferraro 2019: 14; Va. Code Ann. § 18.2-386.2, HB 2678, SB 1736 2019). A törvény értelmében a másról beleegyezése nélkül generált (azaz deepfake-kel készített) képek és videók terjesztését vétségnek minősíti (HB 2678 VA 2019), a vonatkozó szankció az ilyen típusú cselekmény esetében a pénzbüntetéstől akár az egy évig terjedő szabadságvesztésig is terjedhet.

3.2.2. A mesterséges intelligencia szabályozása és az Európai Unió deepfake-felfogása: alacsony kockázat, kevés megkötés

Az európai szabályozásnál elsődlegesen a mesterséges intelligencia szabályozását, az AI Actet (Ződi 2021) kell megemlíteni. A rendeletjavaslat nem ad pontos definíciót a deepfake-re, ugyanakkor a deepfake szabályozását rendezi mint alacsony kockázatú MI-rendszer. Ennek értelmében a deepfake tartalmakra az 52. cikk vonatkozik, amely általános átláthatósági kötelezettségeket ír elő a szolgáltatók részére. Az 52. cikk (3) bekezdése értelmében:

> …azon MI-rendszerek felhasználói, amelyek olyan, meglévő személyekre, tárgyakra, helyekre vagy más szervezetekre vagy eseményekre érzékelhetően hasonlító képet, audio- vagy videotartalmat generálnak vagy manipulálnak, amely egy személy számára megtévesztő módon eredetinek vagy valóságosnak tűnhet („deepfake"), közlik, hogy a tartalmat mesterségesen hozták létre vagy manipulálták.

Ez a szabályozási attitűd, azaz a deepfake mint alacsony kockázatú AI-technológia jogszabályba foglalása igen meglepő és még inkább ellentmondásos, ugyanis nincs logikus magyarázat arra, hogy miért a legminimálisabb átláthatósági szabályok vonatkoznak a fejlesztőkre, miközben a deepfake tartalmakról mind tudományos, mind szakmai körökben megállapításra került, hogy közel 90 százalékuk nem konszenzuális pornográf tartalom (Hao 2021).

4. Jognál jobb szabályozás? Alternatívák és perspektívák a jogon kívül

A probléma valódiságát mi sem mutatja jobban, minthogy 2019 júliusában az ENSZ véleménynyilvánítás szabadságáért felelős különleges jelentéstevője, az Európai Biztonsági és Együttműködési Szervezet (OSCE) sajtószabadság-felelőse, az Amerikai Államok Szervezete (OAS) véleménynyilvánítás szabadságáért felelős különleges jelentéstevője és az Emberek és Népek Jogainak Afrikai Bizottsága (ACHPR) véleménynyilvánítás szabadságáért és az információhoz való hozzáférésért felelős különleges jelentéstevője jubileumi nyilatkozatot adott ki. Ebben vázolták a szólásszabadságra a következő évtizedben váró kihívásokat, és felhívták a világ vezető politikusainak és az online platformok tulajdonosainak a figyelmét az emberi jogok szempontjából érzékeny megoldások elfogadására. Kiemelték a dezinformáció okozta kihívásokkal szemben fellépéseket, beleértve a deepfake egyre gyakoribb megjelenését is (Gosztonyi 2023).

A Facebook 2020-ban bejelentette, hogy betiltja a mesterséges intelligencián alapuló algoritmusokkal készített deepfake videók megosztását (Bickert 2020). Molnár Anita és Rab Árpád (2021) ugyanakkor ezzel kapcsolatban megjegyzi, hogy

> …egy 2020. novemberi – a választási erőfeszítéseket összegző – Twitter-bejegyzés szerint október 27-e óta 300 000 tweethez adták hozzá a félrevezető tartalomra figyelmeztető címkét, ami az ebben az időszakban a választásokkal kapcsolatos összes bejegyzés 0,2 százaléka volt. Deepfake-ről nem tettek említést sem.

Továbbá a Facebook Ellenőrző Bizottság – a cég saját, bíróságszerű testülete (Lendvai 2023b) – először 2023-ban vette napirendre egy deepfake tartalom elbírálását, amely a testület gyűlöletbeszéddel kapcsolatos gyakorlata kapcsán (Lendvai 2023c) meghatározó ítélet lehet.[4]

Az Andrew Ray-féle fentebb említett három problémára három megoldási javaslatot kínálunk. A deepfake tartalmak eltávolítása tekintetében kijelenthetjük, hogy sem az amerikai, sem az európai megoldás nem kínál valós jogi segítségnyújtást. E körben két megoldást tartunk lehetségesnek. Az *ex ante* megoldás lehet a platformokat arra kötelezni, hogy alkalmazzák az egyre fejlődő és a deepfake-demokratizálódás dinamikus evolúcióját lekövető deepfake-felismerő rendszereket (Masood et al. 2022), azaz a videókra a platformok a posztok publikálása után automatikusan rátennének egy jelzést, hogy az adott videó deepfake-technológia felhasználásával készült. Az *ex post* megoldás tekintetében viszont Ray polémiafelvetésével szemben azt gondoljuk, hogy a jog eszköztára nem alkalmas arra, hogy a károk utólagos megtérítésén túl hatékonyan be-, illetve közbeavatkozzon egy választáson induló képviselőjelölt vagy deepfake-dezinformációval érintett személy jóhírnevének visszaállításában – e kérdés nagy hasonlóságot mutat a sajtóhelyreigazítási eljárásokkal szemben megfogalmazott általános kritikákkal.

Javasoljuk viszont, hogy e probléma megoldását a médiaoktatásban és a kritikus gondolkodás elősegítésében keressük: egy politikus szájából „hihetetlennek" hangzó mondat utólagos ellenőrzése ugyanis kézenfekvőbb feloldása lehet a polémiának, mint a jog malmainak lassú őrlésében bízni. Kiemeljük e helyütt a *fact-checker*ek (tényellenőrök) lehetséges szerepét is, akik – akár a platformok alkalmazásában – rövid időn belül ellenőrizni tudnák a deepfake tartalmak eredetiségét. Ray második problémája kapcsán, azaz a dezinformációs fogalmak és jelenségek platformközti harmonizációja tekintetében megemlíthető az Európai Unió 2022-es dezinformációs kódexe (Code of Practice on Disinformation 2022), amely egységes keretrendszert kíván nyújtani a platformok együttműködésével a dezinformációs tevékenységek és a legjobb gyakorlatok kapcsán. A kódex – azon túlmenően, hogy közös konceptuális és gyakorlati lehetőségeket taglal – egyike azon kezdeményezéseknek, amelyekben majdnem minden online óriásplatform aktívan részt vesz. Végül a fogalmi kérdés kapcsán ismét részben elvetjük Ray gondolatait. Bár kétségtelen, hogy a deepfake fogalmi meghatározása komplex feladat elé állítja

4 Érdekesség, hogy a Facebook Ellenőrző Bizottság által tárgyalt eset is egy Joe Bidenről készült „cheapfake", azaz gyenge minőségű deepfake videóról szól (2023-029-FB-UA).

mind a kutatókat, mind a jogalkotót (EPRS 2021), az e tanulmányban említett fogalmak között nagy átfedés mutatkozik, és – Koltay András (2019: 42) gondolatait átformálva – lehetséges, hogy egy ma megfogalmazott és uniformizált fogalom a holnap emberének már egyszerű jogtörténetté válik. Összességében tehát úgy véljük, hogy a fogalmi sokszínűség jóval kevésbé a konceptuális bizonytalanságot, mintsem a deepfake-jelenség gyorsan változó természetét tükrözi.

5. Következtetések: „Nem kell tökéletesnek lennie, csak elég jónak" (Warzel 2018)

A deepfaketartalom-gyártás mint álhírterjesztési eszköz és metódus valószínűsíthetőleg egyre népszerűbb lesz a következendő években (Allen 2022: 77–78), és egyre szélesebb rétegekhez fog eljutni mind fogyasztói, mind készítői oldalról. Charlie Warzel (2018) fenti mondatát kölcsönvéve kijelenthető, hogy a deepfake-problémakör a technológiai fejlődés következtében a közeljövő egyik legizgalmasabb online szabályozási kérdése lesz, hiszen a nem tökéletes, „csak" elég jó hamisított tartalmak is alkalmasak lesznek sok millió ember megtévesztésére.

A fentiekből az is egyértelműnek tűnik, hogy a jog mint egyedüli, partikuláris szabályozó nem elégséges a jogellenes online tartalmakkal szembeni küzdelemben (Herke 2023). Üdvös folyamat a deepfake-érzékelő rendszerek látványos és hatékony fejlődése, illetve helyes irány a nemzeti – sokszor eltérő – szabályozásokon túlívelő nemzetközi egyezmények sora is, ugyanakkor véleményünk szerint elengedhetetlen a társadalmi tudatosítás is. Javasolt egy olyan médiatudatosítási folyamat és oktatás implementálása, amely során az internethasználók a deepfake veszélyein túl azzal is megismerkedhetnének, hogy milyen lépéseket tehetnek a deepfake tartalmak azonosítása kapcsán. Mindez tehát egy olyan holisztikus megközelítést tesz szükségessé, amelyben a „platformszabályozási háromszög" (Gorwa 2019) szereplői közösen tesznek lépéseket a felmerülő problémák orvoslására, amihez nem csupán a jog eszközeit, hanem az általános médiaműveltség (Nagy 2018) fejlesztését is kívánatosnak tartják.

Irodalom

2023-029-FB-UA, https://oversightboard.com/news/698422811785085-oversight-board-announces-two-cases-altered-video-of-president-biden-and-weapons-post-linked-to-sudan-s-conflict/

Aczél Petra (2023): A deepfake mint hazugság: együttműködés a megtévesztésben. In: Aczél Petra & Veszelszki Ágnes (szerk.): *Deepfake: a valótlan valóság,* 30–40. o. Budapest: Gondolat Kiadó.

Allen, Major D. Nicholas (2022): Deepfake Fight: AI-Powered Disinformation and Perfidy Under the Geneva Conventions. *Journal of Emerging Technologies*, vol. 3, no. 2, pp. 2–4, https://doi.org/10.2139/ssrn.3958426

Allyn, Bobby (2022): Deepfake Video of Zelenskyy Could Be 'Tip of the Iceberg' in Info War, Experts Warn. *NPR*, 2022 March 16, https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia/

Bajomi-Lázár Péter (2022): Karaktergyilkosság vagy médiabotrány: mi a különbség? *Médiakutató*, 23. évf. 2. sz. 7–13. o.

Bickert, Monika (2020): Enforcing Against Manipulated Media. *Facebook*, 2020 January 6, https://about.fb.com/news/2020/01/enforcing-against-manipulated-media/

Citron, Danielle K. & Chesney Robert (2019): *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*. Boston: Boston University School of Law.

Cole, Samantha (2017): AI-Assisted Fake Porn Is Here and We're All Fucked. *Vice*, 2017 December 11, https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn/

Cover, Rob (2022): Deepfake Culture: The Emergence of Audio-video Deception as an Object of Social Anxiety and Regulation. *Continuum Journal of Media & Cultural Studies,* vol. 36, no. 4, pp. 609–621, https://doi.org/10.1080/10304312.2022.2084039

Dickinson, Daniel (2018): *Interview with Farnaz Fassihi*. 2018 May 1, https://news.un.org/en/audio/2018/05/1008682/

Dobber, Tom et al. (2020): Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics,* vol. 26, no. 15, pp. 1–23, https://doi.org/10.1177/1940161220944364

Dredge, Stuart (2016): Five of The Best Face Swap Apps. *The Guardian*, 2016 March 17, https://www.theguardian.com/technology/2016/mar/17/five-of-the-best-face-swap-apps/.

European Parliamentary Research Service [EPRS] (2021): *Tackling deepfakes in European policy*.

Ferraro, Matthew F. (2019): Deepfake Legislation: A Nationwide Survey – State and Federal Lawmakers Consider Legislation to Regulate Manipulated Media. *WilmerHale report*, 2019 September 25.

G. Karácsony Gergely (2022): „Ideje megváltoztatni" – Az arcfelismerő rendszerek alkalmazásának alapjogi kockázatai a közösségi média és a deepfake korában. In: Török Bernát & Ződi Zsolt (szerk.): *Az internetes platformok kora,* 299–318. o. Budapest: Ludovika Egyetemi Kiadó.

Gorwa, Robert (2019): The Platform Governance Triangle: Conceptualising the Informal Regulation of Online Content. *Internet Policy Review*, vol. 8, no. 2, https://doi.org/10.14763/2019.2.1407

Gosztonyi Gergely (2023): *Censorship from Plato to Social Media. The Complexity of Social Media's Content Regulation and Moderation Practices*. Cham: Springer Nature Switzerland AG, https://doi.org/10.1007/978-3-031-46529-1

Herke Csongor (2023): Deepfake: áldás vagy átok? *Pro Futuro,* 13 évf. 1. sz., https://doi.org/10.26521/profuturo/2023/1/13334

Hao, Karen (2021): Deepfake Porn is Ruining Women's Lives. Now the Law May Finally Ban it. *MIT Technology Review*, 2021 February 12, https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/.

Ipsos (2022): Internet Users' Trust in the Internet Has Dropped Significantly Since 2019. *Ipsos*, 2022 November 14, https://www.ipsos.com/en/trust-in-the-internet-2022/

Johnson, Ted (2023): Joe Biden Talks about Watching an AI Generated Deepfake of Himself: "I Said, When The Hell Did I Say That?'" *Deadline*, 2023 October 30, https://deadline.com/2023/10/ai-joe-biden-executive-order-1235586979/

Kocsis, Eric (2022): Deepfakes, Shallowfakes, and the Need for a Private Right of Action. *Dickinson Law Review,* vol. 126, no. 2, pp. 621–650.

Koltay András (2019): *Magyar és európai médiajog*. Budapest: Wolters Kluwer.

Lendvai Gergely Ferenc (2023a): Deepfake a szólásszabadság tükrében: Reflexiók a jog perspektívájából. In: Aczél Petra & Veszelszki Ágnes (szerk.): *Deepfake: a valótlan valóság,* 121–138. o. Budapest: Gondolat Kiadó.

Lendvai Gergely Ferenc (2023b): A Facebook „elitbírósága" kritikai megközelítésben: A Facebook Oversight Board nyolc kiemelt kérdése és lehetséges megoldásai. *Iustum Aequum Salutare*, 19. évf. 3. sz. 227–243. o.

Lendvai Gergely Ferenc (2023c): "Pure Rat Country" – Reflections on Case Decision 2022-001-FB-UA of Facebook Oversight Board (Knin Cartoon Case). *The Journal of Digital Technologies and Law,* vol. 1, no. 3, pp. 612–628, https://doi.org/10.21202/jdtl.2023.26

Lima, Cristiano (2023): A Fake Biden Video Shows the Limits of Meta's Deepfake Policies. *The Washington Post*, 2023 October 13, https://www.washingtonpost.com/politics/2023/10/10/fake-biden-video-shows-limits-meta-deepfake-policies/

Masood, Momina, Marriam Nawaz, Khalid Mahmood Malik, Ali Javed & Aun Irtaza (2022): Deepfakes Generation and Detection: State-of-the-art, Open Challenges, Countermeasures, and Way Forward. *Applied Intelligence,* vol. 53, no. 4, pp. 1–53, https://doi.org/10.1007/s10489-022-03766-z

Molnár Anita & Rab Árpád (2021): A deepfake technológia és az amerikai elnökválasztás. *ITKI Blog*, 2021. I. 18., https://www.ludovika.hu/blogok/itkiblog/2021/01/18/a-deepfake-technologia-es-az-amerikai-elnokvalasztas/.

Mráz Attila (2021): Deepfake, demokrácia, kampány, szólásszabadság. In: Török Bernát & Ződi Zsolt (szerk.): *A mesterséges intelligencia szabályozási kihívásai,* 249–277. o. Budapest: Ludovika Egyetemi Kiadó.

Nagy Krisztina (2018): *Műveltség – Média – Szabályozás. A médiaműveltség médiapolitikai jelentősége és szabályozási keretei*. Budapest: Gondolat Kiadó.

Pantserev, Konstantin A. (2020): The Malicious Use of AI-Based Deepfake Technology as the New Threat to Psychological Security and Political Stability. In: Hamid Jahankhani et al. (eds.): *Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity*, pp. 47–48. Cham: Springer Nature Switzerland AG, https://doi.org/10.1007/978-3-030-35746-7_3

Rankin, Jennifer (2018): Viktor Orbán: Re-election of Hungary's Anti-immigrant Leader is Major Challenge for EU. *The Guardian*, 2018 April 9, https://www.theguardian.com/world/2018/apr/09/viktor-orban-re-election-hungarys-anti-immigrant-leader-major-challenge-for-eu/

Ray, Andrew (2021): Disinformation, Deepfakes and Democracies: The Need for Legislative Reform. *UNSW Law Journal*, vol. 44, no. 3, pp. 983–1013, https://doi.org/10.53637/DELS2700

Reid, Shannon (2021): The Deepfake Dilemma: Reconciling Privacy and First Amendment Protections. *Journal of Constitutional Law,* vol. 23, no. 209, pp. 209–238.

Satarino, Adam & Paul Mozur (2023): The People on Screen are Fake. The Disinformation is Real. *The New York Times*, 2023 February 7, https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html.

Sorbán Kinga (2020): A bosszúpornó és deepfake pornográfia büntetőjogi fenyegetettségének szükségességéről. *Belügyi Szemle*, 68. évf. 10. sz. 81–104. o., https://doi.org/10.38146/BSZ.2020.10.4

Spellberg, Claire (2019): Jim Carrey Transforms into Jack Torrance in 'The Shining' Deepfake Series. *Decider*, 2019 July 11, https://decider.com/2019/07/11/jim-carrey-the-shining-deepfakes/.

Szűts Zoltán (2018): *Online*. Budapest: Wolters Kluwer.

Vaccari, Cristian & Andrew Chadwick (2020): Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society,* vol. 6, no. 1, https://doi.org/10.1177/2056305120903408

Van der Sloot, Bart & Yvette Wagensveld (2022): Deepfakes: Regulatory Challenges for the Synthetic Society. *Computer Law & Security Review,* vol. 46, no. 105716, pp. 1–15, https://doi.org/10.1016/j.clsr.2022.105716

Veszelszki Ágnes (2022): A tudományos influencerektől a deepfake-ig. A legújabb tudománykommunikációs lehetőségek. *Filológia,* 13. évf. 1–4. sz. 27–39. o.

Vincent, James (2017): AI Tools will Make it Easy to Create Fake Porn of Just About Anybody. *The Verge*, 2017 December 12, https://www.theverge.com/2017/12/12/16766596/ai-fake-porn-celebrities-machine-learning/.

Warzel, Charlie (2018): Believable: The Terrifying Future Of Fake News. *BuzzFeedNews*, 2018 February 12, https://www.buzzfeednews.com/article/charliewarzel/the-terrifying-future-of-fake-news/.

Yahoo UK (2023): 'When the hell did I say that?' – Biden Talks about Deepfake Video of Himself. *Yahoo News UK*, 2023 October 31, https://uk.news.yahoo.com/hell-did-biden-talks-deepfake-064430768.html.

Ződi Zsolt (2021): Az Európai Bizottság Mesterséges Intelligencia Kódexének tervezete. *Gazdaság és Jog,* 26. évf. 5. sz. 1–3. o.

Jogszabályok

H.R. 3230 § 1041, https://www.congress.gov/bill/116th-congress/house-bill/3230

H.R. 3230 § 1041(n)(3), https://www.congress.gov/bill/116th-congress/house-bill/3230

Californian Elections Code §20010, https://ocvote.gov/apps/legtracker/elections-code/contents/

Va. Code Ann. § 18.2-386.2, https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.2/

HB 2678, SB 1736 2019, https://lis.virginia.gov/cgi-bin/legp604.exe?191+sum+HB2678

HB 2678 VA 2019, https://lis.virginia.gov/cgi-bin/legp604.exe?191+sum+SB1736

Code of Practice on Disinformation 2022, https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation