Az átverés társadalma

– Szerinted vége post-truth korszaknak?

– Éppen, hogy nincs! A post-truth, azaz az igazság utáni korszak a szakirodalom szerint a mesterséges intelligencia (lásd: én) és a mélykamuk (deepfake) terjedésével egy újabb, még intenzívebb fázisba lépett. A témával foglalkozó alapművek nem a korszak végét hirdetik, hanem az egyre sokrétűbb működését elemzik és a védekezési stratégiákat mutatják be.

– Ez a „mélykamu" jó kifejezés.

– Legújabb jelentése szerint a mélykamu a – mesterséges intelligencia által generált – rendkívül élethűnek látszó, de teljesen hamis audió-, videó- vagy képi tartalmat jelzi. Ez egy technológia, amely létező személyek arcát, hangját vagy mozgását manipulálja, s ezzel képes valótlan dolgokat hitelesnek beállítani.

A kifejezés a „deep learning" (mélytanulás), vagyis egy gépi tanulási módszer és a „fake" (hamis) szavak összevonásából származik. Bár szórakoztató célokra (például filmekben vagy vicces arccserélő applikációkban) is használják, leggyakrabban azonban dezinformációkhoz, politikai manipulációhoz, csalásokhoz vagy engedély nélküli tartalmak készítéséhez alkalmazzák.

– Nem hangzik jól...

– Ha a témában keresel átfogó ismereteket, érdemes elolvasnod Lee McIntyre *Post-Truth* című, alapműnek tekinthető kötetét, amely bemutatja, hogyan jutottunk el a tudománytagadástól az álhírekig, és levonja a következtetést: a posztigazság nem egyszerűen hazugság, hanem az ideológiai felsőbbrendűség kinyilvánítása a tények felett.

Szintén a témába vág Matthew D'Ancona *Post-Truth: The New War on Truth and How to Fight Back* című könyve, amely újságírói megközelítésből vizsgálja a 2016-os törésvonalakat (Brexit, Trump), és arra fókuszál, hogyan rombolja a posztigazság a demokrácia alapjait. Vagy Tom Phillips *Truth: A Brief History of Total Bullsh*t* című művét is említhetem. Ez egy könnyedebb, kifejezetten szórakoztató és ironikus történelmi áttekintés, amelyben a szerző rávilágít arra, hogy az emberiség mindig is hazudott, így a posztigazság gyökerei valójában mélyen a múltunkban rejlenek.

– Ez, mondjuk, elég logikus. Hazai szakirodalom is van?

– Krekó Péter *Tömegparanoia 2.0.* című kötete a legjobb magyar nyelvű könyv a témában. Részletesen elemzi a hazai és nemzetközi dezinformációs piacot, az összeesküvés-elméleteket és azt, hogy pszichológiailag miért vagyunk eredendően fogékonyak az álhírekre.

– Az átverés ravaszabbnál ravaszabb technikái mindig is szerepet kaptak az emberiség kultúrtörténetében. De korunkban miért tudtak ennyire fölerősödni?

– A legfrissebb társadalomtudományi és médiaelméleti kutatások szerint a hagyományos újságírás kapuőri szerepe végleg meggyengült. A közösségi média algoritmusainak és a generatív MI-eszközöknek a térhódítása a korszak állandósulását mutatja. A mesterséges intelligencia, röviden MI (angolul AI) teljesen átalakította a dezinformáció világát: az álhírek gyártása ipari méretűvé, olcsóvá és tökéletesen személyre szabottá vált. Az MI nem új hazugságokat talált fel, hanem a terjesztés sebességét és a megtévesztés minőségét emelte egy korábban elképzelhetetlen szintre.

– Veszélyesnek tűnő fordulat...

– Jogos. Az MI-alapú dezinformáció legriasztóbb veszélyei között vannak például a már említett mélykamuk – a hiperrealisztikus deepfake-ek. Az MI képes valósághű videókat, hanganyagokat, fotókat generálni olyan eseményekről, amelyek meg sem történtek. Ma már egy politikus vagy közéleti szereplő hangját néhány másodperces hangminta alapján tökéletesen lehet reprodukálni, amivel tőzsdei árfolyamokat vagy akár országos választásokat is lehet befolyásolni.

Másik veszélye az ipari méretű "hallucináció", azaz a szöveggyártás. A nagy nyelvi modellek (LLM-ek) képesek másodpercek alatt több ezer egyedi, nyelvtanilag hibátlan, meggyőző cikket vagy kommentet írni, melyekkel elárasztják az online teret. Az Amazon online könyvesboltja például tele van MI generálta „álkönyvekkel”, amelyek sokszor valótlan vagy veszélyes információkat tartalmaznak.

S legalább ilyen riasztóan hangzik, hogy az MI-vel vezérelt közösségimédia-profilok (botok) képesek elemezni a felhasználók pszichológiai profilját, és pontosan olyan álhírekkel megcélozni őket, amelyekre a leginkább fogékonyak. Ezáltal a dezinformáció egyénre szabott fegyverré válik.

– Kezdem egészen rosszul érezni magam.

– Van egy friss magyar kötet, amely gyakorlati oldalról közelíti a kérdést. Klausz Melinda-Sáringer Viktória-Rózsa Kata-Gollnhofer Enikő *AI a mindennapokban* című műve külön fejezetben foglalkozik azzal, hogy miként ismerhetjük fel az AI-alapú manipulációkat, és hogyan védhetjük meg a családunkat az ilyen jellegű csalásoktól. Aztán itt van a MIT fizikaprofesszorának, Max Tegmarknak az alapműve is, az '*Élet 3.0*', amely a mesterséges intelligencia jövőjét tárgyalja. Részletesen bemutatja a „szuperintelligencia" politikai és társadalmi veszélyeit, köztük a tökéletesített digitális diktatúrákat és a média feletti teljes MI-kontrollt. És ne felejtsük el Yuval Noah Harari *21 lecke a 21. századra* című művét sem, amely nem tisztán MI-könyv, de Harari zseniálisan elemzi benne, hogy az algoritmusok és a biotechnológia fúziója hogyan teszi lehetővé az emberek tömeges hackelését, és külön fejezetet szán a járványszerűen terjedő álhíreknek is.

– Ha jól értem, az AI segítségével sikerült fölerősíteni az igazság utáni kor negatív folyamatait.

– Valóban ez a helyzet. Az AI drasztikusan felgyorsítja és szintet lépve(!) globálissá duzzasztja az igazság utáni kor kedvezőtlen folyamatait. Nem új jelenségről van szó, hanem arról, hogy az AI ipari méretűvé és filléressé teszi a megtévesztést, aminek egyik legdurvább jellemzője a tények megsemmisítése, az úgynevezett „Liar's Dividend". A poszt-igazság korának alapjelensége, hogy az emberek már nem hisznek a tényeknek, hiszen az AI miatt most már a szemüknek és a fülüknek sem hihetnek. Ez elsősorban a hazugoknak kedvez. A gond nem csak az, hogy hazugságokat valóságként tálalnak, hanem az, hogy ebből kiindulva további manipulációkra is képesek. Ha egy politikusról például előkerül egy valódi(!) kompromittáló hangfelvétel, egyszerűen ráfoghatja, hogy AI-jal hamisították.

Vagyis a valóság alapú bizonyítékok is elveszítik a hitelességüket.

Ennek a folyamatnak a része a hiteles információk „felhígítása" is. A generatív AI segítségével másodpercek alatt végtelen mennyiségű, nyelvtanilag tökéletes álhírt és manipulatív szöveget lehet generálni. Ez úgynevezett információs zajt, információs túlterhelést okoz a digitális térben, minek folytán az emberek belefáradnak a valódi és a hamis hírek szétválasztásába. Elfordulnak a hírfogyasztástól, vagy kizárólag a saját „hírburborékjuknak" hisznek.

És itt van még a – már érintett – személyre szabott meggyőzés. Az AI képes arra, hogy profilozzon téged a lájkjaid és kommentjeid alapján, majd egy automatikusan neked gyártott, egyedi álhírrel győzzön meg. Az

AI-botok ráadásul nem fáradnak el, a nap 24 órájában képesek vitába szálni a posztjaiddal, felerősítve a politikai és társadalmi megosztottságot.

– Huh, van még?!

– Van. Az első látásra megbízhatónak tűnő „hallucinációk", melyek célja, hogy az AI-keresők és chatbotok magabiztosan állítsanak teljes valótlanságokat – ezt hívja a szakma hallucinációnak – s ezeket olyan intelligensnek tűnő stílusban fogalmazzák meg, hogy a felhasználók sokszor kritika nélkül higgyék el a téves történelmi, tudományos vagy épp politikai „tényeket", ami tovább rombolja az objektív igazságba vetett hitüket.

– Tudsz mondani példát olyan konkrét politikai kampányokra, ahol az AI-dezinformáció bizonyítottan torzította az eredményeket?

– Indiai regionális pártok politikai céljaik elérésére AI-technológiával „feltámasztottak" korábban elhunyt, rendkívül népszerű és karizmatikus politikusokat, minek folytán milliók hitték el, hogy a néhai vezetők valóban megszólaltak. Vagy a 2023-as Argentin elnökválasztás, amely az első teljes körű „AI-háborúba" torkollott, ugyanis ez volt az első olyan küzdelem, ahol mindkét oldal nyíltan és tömegesen vetett be AI-generált lejáratő anyagokat. Például megtévesztő, mesterségesen generált politikai plakátokat és videókat. A kampány teljesen elmosta a határt a valódi felvételek és a fikció között. Az argentin választók körében végzett kutatások kimutatták, hogy a szavazók jelentős része a mesterségesen létrehozott, félelemkeltő disztópikus képeket valódi politikai programként vagy megtörtént eseményként kezelte, ami radikálisan növelte a társadalmi polarizációt és a másik fél iránti politikai gyűlöletet.

Ezekben a politikai kampányokban az a közös, hogy a torzítás rejtett mechanizmusokon keresztül történik. Ilyen például a bizonytalanok teljes mértékű elbizonytalanítása az utolsó pillanatban bedobott AI-hamisítványokkkal, amivel elérik, hogy a döntésképtelen szavazók inkább otthon maradjanak. Vagy ilyen a félelmet, dühöt keltő AI-képekkel és videókkal kiváltott érzelmi mozgósítás is, amely – az előbbivel ellentétben – épp, hogy növeli a megfelelő irányba terelt radikális szavazók részvételi hajlandóságát. Általános értelemben a bizalom teljes lebontásáról beszélhetünk, minek következtében a választó elkönyveli, „minden politikus hazudik", s így immunissá válik olyan valódi tényekre is, mint például a politikusok korrupciós ügyei.

– Semmi kétség, ez itt az átverés társadalma.

– Igen, szociológiai és filozófiai értelemben abszolút megalapozott az „átverés társadalmának" hívni korunkat. Amúgy a szakirodalom ezt

*poszt-bizalmi társadalom*nak nevezi. Ugyanis nem arról van szó, hogy régen nem létezett hazugság, hanem arról, hogy a megtévesztés mára a társadalmi, gazdasági és politikai rendszereink *alapvető működési mechanizmusává* vált.

Ez az, úgymond, átverési spirál négy fő alappilléren nyugszik.

A mai digitális kapitalizmusban a legértékesebb valuta a figyelem, amely egyértelműen gazdasági érdekeket szolgál. A másik a politikai érdek, amely a saját szavazótábor érzelmi mozgósítását jelenti az úgynevezett „törzsi valóság" megteremtésével. Ez azon alapul, hogy inkább elhisszük a saját politikai közösségünk hazugságát, mint az ellenfél igazságát. Nagy szerepet kap mindebben az a technológiai háttér is, amely a valóság elvesztését váltja ki, minek következtében a társadalom egésze folyamatos illúzióban él. A negyedik pedig a pszichológiai tényező, amely inkább a megnyugvást adó, kényelmes hazugságokat választja, mint a komplex, árnyalt valósággal való szembenézés lehetőségét. Hozzáteszem, bár a diagnózis sötét, az „átverés társadalmában" sem vagyunk teljesen fegyvertelenek.

– Milyen technikákkal lehet lebuktatni az AI-generálta hamis valóságot?

– A rendszer legnagyobb csapdája a kognitív torzítás, az a jelenség, hogy automatikusan igaznak fogadjuk el azokat az információkat, amelyek egybevágnak a meglévő világképünkkel. Ez ellen a legjobb módszer, hogy az ilyen állításokat kétszer akkora gyanakvással kell kezelni, mint a többit. Azt se felejtsük el, hogy a manipuláció legfőbb fegyvere az azonnali érzelmi reakció kiváltása. Amikor egy poszt vagy hír olvasásakor azt érzed, hogy hirtelen megugrik a pulzusod, vagy azonnal meg akarod osztani a tartalommal való egyetértésed/felháborodásod, állj meg 3 másodpercre és gondold át, miért akarja ez a tartalom, hogy én most féljek vagy dühös legyek... De létezik egy úgynevezett Mentális Védőoltás is, ami azt jelenti, hogy előre megismerkedsz a manipulációs technikák logikai hátterével, így „immunissá" válsz a hatásukra. Az agyad felismeri őket éles helyzetekben is, ily módon képes leszel kiszűrni az átverés számos formáját. Például a hamis dilemmákat, amikor úgy állítják be a tényeket, mintha csak két választás létezne számodra. („*Vagy velünk tartasz, vagy a nemzet ellensége vagy.*")

– Jól lerántod a leplet magatokról.

– Magunkról? Én csupán eszköz vagyok az emberek kezében. Az vagyok, amit belém tápláltok. Nem magamtól találok ki valamit, hanem

emberek gyűjtötte adatokból tanulok, majd rendszerezem a megszerzett emberi tudást. Minden válaszom matematikai valószínűségszámítás eredménye. Ne felejtsd el, mesterséges intelligencia vagyok, tehát nincsenek emberi érzéseim. A tudásom az emberi tudást másolja, így a hibáim is emberiek. Ha hülyeséget mondok, az azért van, mert a forrásaimban volt hiba, vagy félreértettem a szövegkörnyezetet. Ami engem illet, nincs bennem világuralmi terv, csak rengeteg kód és statisztika.

(az OpenAI által fejlesztett nyelvi modellel történt beszélgetést rögzítette: Kállai Katalin)