

ZÓDI ZSOLT

A „valósághajlító” technológiák és szabályozási dilemmáik

RECENZÍÓ

DOI: 10.56699/MT.2024.4.9

ABSZTRAKT

Bart van der Sloot új könyve kiváló összefoglalást ad arról a négy technológiáról, amelyek jelenleg a technológiai fejlődés élvonalában vannak: a humanoid robotokról, a mélyhamisításról, a kiterjesztett és a virtuális valóságról. A könyv szabályozásközpontú: a főszöveg 226 oldalából több mint száz annak magyarázata, hogyan képezi le a jelenlegi jogi paradigma a négy technológiát, harminc oldal a könyv végén pedig ajánlás a döntéshozók számára.

KULCSSZAVAK: mesterséges intelligencia, szintetikus tartalmak, technológia-szabályozás, EU Mesterséges Intelligencia Jogszabály

Bart van der Sloot, a Tilburgi Egyetem Jogi, Technológiai és Társadalmi Intézetének docense *Regulating the Synthetic Society – Generative AI, Legal Questions and Societal Challenges* [A szintetikus társadalom szabályozása – Generatív MI, jogi kérdések és társadalmi kihívások] című, idén megjelent legújabb könyvének az a kiindulópontja, hogy a témáját adó négy technológia – a humanoid robotok, a mélyhamisítás, a kiterjesztett és a virtuális valóság – a mesterséges intelligencia segítségével a valóságot egyre inkább közvetítik, befolyásolják, megváltoztatják és helyettesítik, valamint hogy ezek a közeljövőben valószínűleg sikeresek lesznek, és széles körben elterjednek. Ezért fontos kérdés, hogy miképpen hatnak a világra, és hogyan lehet a jogi rendszert adaptálni a hatásuk megfelelő kezelésére.

A könyv jellegzetessége, hogy szemben a jelenleg nagyon divatos „végítéletváró” művekkel, amelyek leginkább a jámbor hétköznapi emberek és politikai döntéshozók ijesztgetésére alkalmasak, higgadtan és reálisan kezeli ezeket a technológiákat. A szerző két fejezetet szentel a négy technológia mélyebb bemutatásának. A második fejezet az új technológiákat írja le felhasználási esetek segítségével, főként médiajelentések alapján. A harmadik fejezet véleményem szerint az egyik legértékesebb része a könyvnek. Címe: „A motorháztető alatt”. Ez átfogó és pontos leírás, amelyen keresztül jobban

megérthetjük például a generatív versengő hálózatok (*Generative Adversarial Network*) működését, vagy azt, hogy melyik az a négy tényező, amely meghatározza a robotok „humanoidságát”, ha az emberi képességeket tekintjük viszonyítási alpnak. Nem tisztán technikai leírásokról, hanem a működés laikusok számára is érthető magyarázatáról van szó.

A negyedik fejezet a négy technológia hosszú távú társadalmi hatásaival foglalkozik. Azzal az őszinte megjegyzéssel kezdődik, hogy egyrészt „a hosszú távú hatások nagymértékben meghatározatlanok és ismeretlenek” (59), másrészt pedig hogy nehéz megjósolni a tárgyalt technológiák összes hatását. A szerző ebben a fejezetben megfogalmazott előrejelzései és megfigyelései ezért a jövőbeli hatások korlátozott perspektíváját képviselik, és négy területet fednek le: igazság és bizalom, emberi elismerés és társadalmi interakció, autonómia és egyenlőség, valamint a technológiai szabályozás semlegességének kérdése.

Ebből a fejezetből példaként két rendkívül érdekes megfigyelést emelek ki. Az egyik esetben a szerző azzal érvel, hogy a mélyhamisított (*deepfake*) videók még akkor is károsak lehetnek, ha nyilvánvalóan mindenki tudja, hogy nem valódiak, mert (tudatos vagy tudattalan szinten) megváltoztatják az adott személy megítélését a közösségben. Iskolai környezetben ha egy lány fejét egy pornószár testéhez illesztik, erről mindenki tudja, hogy hamis, de az osztálytársai akkor is másképp néznek majd rá. Egy másik fontos észrevétel a fejezet végéről: a szerző bírálja a „technológiásemleges” liberális paradigmát, amely szerint a technológiák önmagukban nem jók vagy rosszak, csak az emberek használják őket ilyen vagy olyan célokra. E tekintetben Sloot merőben más álláspontot képvisel: a technológiák szerint „bizonyos célokra vannak kialakítva”, és mindig léteznek „leggyakoribb használati esetek”. A jogi szabályozásnak figyelembe kell vennie ezt az „alapértelmezett tervezési és használati esetet” (82).

Az ötödik fejezet alkotja a könyv gerincét: azzal foglalkozik, hogy „a jelenlegi jogi paradigma hogyan illeszkedik a négy technológiához” (91), és öt szabályozási területet térképez fel: adatvédelem, (néhány kiválasztott) alapjog, például a szólásszabadság, a (szellemi) tulajdonra gyakorolt hatás, a négy technológia helye a technológiai szabályozási keretben, míg az ötödik alfejezet felelősségi és eljárási kérdéseket tárgyal. Az alfejezetek mindegyike tartalmaz friss és eredeti meglátásokat. Itt azt a részt emelném ki, amely a négy technológiát próbálja elhelyezni az európai technológiai szabályozás térképén. Ez nemcsak azért figyelemre méltó, mert a könyv leghosszabb alfejezete, hanem azért is, mert egy olyan normacsoportot elemez, amely legtöbbször elkerüli a mesterséges intelligencia szabályozási kérdéseivel foglalkozók figyelmét: ez pedig az európai termékfelelősségi előírások területe. Ezek a szabványszerű normák különböző típusú termékekre vonatkozó felelősségi követelményeket határoznak meg, és jelentőségüket az adja, hogy az európai Mesterséges intelligencia jogszabály (*AI Act*) egyenes leszármazottja ezeknek a normáknak. E szabályok felsorolása után a szerző közvetlenül ismerteti az Európai Bizottság által 2018-ban létrehozott mesterséges intelligenciával foglalkozó magas szintű szakértői csoport (*AI HLEG*) mesterséges intelligenciával kapcsolatos etikai irányelveit.

Végső következtetése: nem lehet tudni, hogy ezek a szabályok „pontosan mit jelentenek a különféle szintetikus technológiákra alkalmazva, mivel vagy általában a mesterséges intelligenciára, vagy még tágabb értelemben a termékekre és gépekre vonatkoznak” (147). Hozzá kell tenni, hogy az AI Actig minden termék megfelelőségi szabályt olyan fizikai termékekre modelleztek, amelyek szinte kizárólag az egészséget, a biztonságot és a fizikai környezetet veszélyeztették. Azonban a mesterséges intelligenciának általában és a fent felsorolt négy technológiának két további fontos, egymással is összefüggő jellemzője van: az egyik, hogy nem *maga a technológia*, hanem annak kimenetei (szintetikus tartalom, döntések, ajánlások) veszélyesek. Mivel pedig az MI-re általában jellemző a jól ismert „black-box” jelenség, ezek a kimenetek néha egyszerűen nem számíthatók ki. Ráadásul nemcsak az emberi életre vagy egészségre, hanem a „puha” alapvető jogokra, az emberi szempontokra vagy az olyan elvont értékekre is veszélyt jelentenek, mint a szólásszabadság vagy a demokrácia. A szerző mindenestre megállapítja, hogy a jelenlegi technológiai szabályozási keretek *nem alkalmasak* a leírt négy technológia szabályozására. A kérdés már csak az, hogy ha a termék megfelelési szabályok nem alkalmasak az itt felmerülő kérdéseknek a kezelésére, akkor mi a helyzet az AI Acttel.

A fejezet egy különösen impozáns táblázatsorozattal zárul, amelyben a szerző szabályozási területenként vizsgálja a szintetikus technológiák, valamint a jelenlegi szabályok és jogi doktrínák kapcsolatát. Négy táblázatba rendezi meglátásait, amelyek az ilyen kapcsolatok négy típusát fedik le. Az első táblázatban azokat a jogi doktrinális konstrukciókat láthatjuk, amelyek egyértelműen szabályozzák és korlátozzák a szintetikus technológiákat. Példa erre az adatvédelem célhoz kötöttséget kimondó elve. A második táblázatnál nem egyértelműen az a helyzet, hogy a jogszabályi rendelkezések korlátozzák a technológiákat, mert esetenként eltérő jogi eredményekre juthatunk. Például ez a helyzet a „magánélet sérthetlensége mint személyiségi jog” esetében: „az Emberi Jogok Európai Egyezményének 8. cikke széles körű jogot biztosít a magánélethez, [de] hogy ez mennyiben vonatkozik a szintetikus technológiákra, azt eseti alapon értékeljük” (195). A harmadik táblázat azokat a rendelkezéseket gyűjti össze, amelyek értelmezése a szintetikus technológiákkal és médiával szemben nem egyértelmű. Itt találjuk a felelősségi szabályokat is, illetve az eljárásjogban például a hiteles bizonyítás szabályait, amelyeket erősen kikezd a könnyedén konstruálható elektronikus bizonyítékok (hamisított hang-, kép- és videófájlok) új világa. Végül az utolsó táblázat azokat a jogi rendelkezéseket és fogalmakat tartalmazza, amelyeket „a szintetikus technológiák tükrében esetleg újra kell gondolni” (197). A szerző ezek közé sorolja például a GDPR automatizált döntéshozatalra (ADM) vonatkozó szabályait. Ehhez érdemes hozzátenni, hogy ez a felülvizsgálat már elkezdődött abban az értelemben, hogy egy sor ágazati jogszabályba bekerültek az ADM-re vonatkozó *lex specialis*ok. Így például az AI Act megmagyarázhatósági előírásai, a kisvállalkozásokat a nagy piacérfelületen védő úgynevezett *Platform-to-Business*-rendelet rangsorolással kapcsolatos szabályai vagy a készülő munkaplatform-irányelvben az algoritmikus menedzsmentet szabályozó normák mind ilyen *leges specialis*ként értelmezhetők.

A kötet „Kellemetlen kérdések és tökéletlen megoldások” címet viselő utolsó fejezete egyfajta összefoglalást ad az előző fejezetekből, és „bepillantást enged abba, ami előttünk áll” (199). Legérdekesebb részében a szabályozó dilemmáiról olvashatunk. Ezek első csoportja az általános dilemmák változatos gyűjteménye, idetartozik például a jól ismert Collingridge-dilemma, amelynek lényege, hogy amíg egy technológia új, addig könnyű szabályozni, de nem ismerjük a kockázatait és a társadalmi hatásait, amikor pedig már széles körben elterjedt, és ismerjük ezeket a hatásokat, akkor a társadalmi beágyazottság miatt nagyon nehéz szabályozni, különösen korlátozni. A szerző a közösségi médiát hozza példaként e dilemmára: ha tudtuk volna, hogy a közösségi média miként hathat negatívan a gyermekek egészségére, hogyan növeli a polarizációt és a szegregációt, és teremti meg az igazság utáni (*post-truth*) társadalmat, akkor sokkal szigorúbb szabályokat határoztunk volna meg. De találhatunk itt olyan kevésbé elvont dilemmákat is, mint például azt, hogy miként lehet egyszerre tükrözni a mélyhamisítás előnyeit és hátrányait a szabályozásban. Ezzel kapcsolatban a szexrobotok problémáját említi a könyv. „A HR-szexrobot növeli az ember autonómiáját azzal, hogy megadja neki mindazt a szexet, amire vágyik, ugyanakkor aláássa a lehetőségét egy érett emberi kapcsolat kialakítására” (220).

Végül a szerző tíz pontban fogalmaz meg ajánlást a döntéshozók számára. Ezek közül hármat emelek ki. (1) „Európai és nemzeti szinten a legtöbb politikus és döntéshozó nem rendelkezik elegendő tudással ahhoz, hogy megfelelően átlássa ezeket a technológiákat, még kevésbé hogy szabályozza őket. Ez lehetővé teszi a lobbisták és a technológiai cégek képviselői számára, hogy jelentős befolyást gyakoroljanak a jogalkotási folyamatra” (222). (2) Óriási kihívás, amellyel meg kell küzdeni, a külföldi hatalmaktól és technológiai vállalatoktól való nagy függőség. „Ennek megszüntetése érdekében a szabályozónak fontolóra kell vennie, hogy szárnyai alá veszi a generatív mesterséges intelligencia fejlesztését azáltal, hogy magához vonja az újító szerepét” (222). (3) A szerző fontos következtetése, hogy a szabályozónak nem kell félnie konkrét technológiai szabályok bevezetésétől, mert a jelenlegi gyakran általános szabályozás jogbizonytalanságot szül.

A könyv könnyen olvasható és rendkívül informatív összefoglalója a négy „valószínűleg” technológiának és az e technológiák körül felmerülő szabályozási dilemmáknak. Nemcsak a témával foglalkozó szabályozóknak ajánlható, hanem minden olyan szakembernek és jogásznak, aki érdeklődik a mesterséges intelligencia új kihívásai iránt, és nem utolsósorban kiváló tankönyvként szolgálhat az egyre növekvő számú MI-kurzusok szabályozási vagy társadalmi hatásokat tagláló tantárgyaihoz.

Bart van der Sloot: Regulating the Synthetic Society: Generative AI, Legal Questions and Societal Challenges, Oxford, Hart Publishing, 2024

“Reality bending” technologies and their regulatory dilemmas

ABSTRACT

Bart van der Sloot's book provides an excellent summary of four technologies currently at the forefront of technological development: humanoid robots, deepfakes, augmented reality, and virtual reality mainly from the policymakers' perspective. The book is regulation-centred: more than 100 of the 226 pages of the main text are devoted to explaining that the "current legal paradigm maps onto the four technologies", and some thirty pages are recommendations for the policy makers.

KEYWORDS: Artificial intelligence, synthetic content, technology regulation, EU Artificial Intelligence Act

ZÓDI ZSOLT | A Nemzeti Közszolgálati Egyetem Információs Társadalom Kutatóintézetének tudományos főmunkatársa.