

BALOGH ZSOLT GYÖRGY

AZ MI-RENDSZEREK ELLENŐRZÉSE, FELÜGYELETE ÉS MONITOROZÁSA

BEVEZETÉS

Az emberiség életmódját alapvetően átalakították az információs korszak újításai és az ezeket tömegesen elérhetővé tevő multinacionális technológiai óriáscégek. Miközben a fizikai világban az éghajlati viszonyok drámaian megváltoznak, pusztító háborúk törnek ki, valamint egy vírusos világjárvány is próbára tette a civilizációt, a virtualitás beszivárog a mindennapi életbe. Az emberek a mobil eszközök képernyőjéhez tapadva töltik az idejük számottevő hányadát, és reménytelen pszichológiai csapdában vergődve frissítik újra és újra a személyre szabott tartalmak végeláthatatlan folyamát.

Ez csak egy a mesterséges intelligencia (MI) számos alkalmazása és az életünkre gyakorolt hatása közül. A közösségi hálózatok kifinomult, mentális trükköket alkalmazó algoritmusai nagyrészt a mesterséges intelligenciát hasznosító fejlett profilalkotási képességeken alapulnak. Az elmúlt évtizedek során a *mesterséges intelligencia* az informatika egyik legfejlettebb, legmesszebbre mutató és legnagyobb kihívást jelentő területévé fejlődött. Az alkalmazások széles és gazdagodó köre bontakozik ki előttünk, és ezzel párhuzamosan egyre több problémával szembesülünk az MI felhasználását övező etikai, jogi és irányítási kérdésekkel kapcsolatban.

Mielőtt érdemben tárgyalnánk e technológia jogi és felelősségi kérdéseit, ismertetni kell a mesterséges intelligencia tulajdonságait, képességeit és funkcióit. Következésképpen e jogi áttekintés céljából elkerülhetetlen lesz az MI-rendszerek meghatározása és taxonómiája.

DEFINÍCIÓ ÉS RENDSZEREZÉS

A közelmúltban elterjedt, technológiával kapcsolatos szakirodalom rengeteg definíciót és megközelítést kínál a mesterséges intelligenciáról önmagában és a működő, megvalósított mesterségesintelligencia-rendszerekről. A jogi diskurzus kedvéért most csak néhány – talán önkényesen kiválasztott – megközelítést veszünk figyelembe.

A mesterséges intelligencia tudományos definíciói

A mesterséges intelligenciára vonatkozóan nincs teljes közmegegyezésnek örvendő, végleges definíció, több értelmezés egyes elemeit is figyelembe vehetjük. Az intelligencia fogalma több attribútummal azonosítható és mérhető.

Eszerint az intelligencia próbaköve lehet:

- az emberhez hasonló kognitív teljesítmény vagy
- egy elvont, ideális racionalitás, azaz a helyes gondolkodás.

Más meghatározások szerint az intelligencia létének és érettségének mércéje a kifinomult:

- gondolkodási folyamat és érvelés képessége vagy
- az intelligens viselkedés lenne.

Ezek a szempontok Russell és Norvig¹ leírása szerint a következőképpen határozzák meg a mesterséges intelligencia négy fő értelmezését:

- *emberi* módon történő *gondolkodás*, azaz kognitív modellezés;
- *emberi* lényként történő *viselkedés*, azaz képes megfelelni a Turing-teszten;
- *racionális gondolkodás*, azaz logikai alapú modellezés;
- *racionális cselekvés*, azaz racionáliságens-modell.

Russell és Norvig egyúttal hangsúlyozza, hogy az *intelligens ágens*² a mesterséges intelligencia fent említett kategóriáinak központi fogalma. Az intelligens ágens arra szolgál, hogy a környezetből észleléseket fogadjon és cselekvéseket hajtson végre. A híres Turing-teszt, amelyet maga Turing „*imitációs játéknak*”³ nevezett, és egyfajta próbaköve a gép intelligens viselkedésének, néhány releváns kritériumot állít fel. Ennek ellenére bizonyos vélemények szerint egy *Eugene Goostman* nevű chatbot 2014-ben átmert a Turing-teszten, bár ezt a szenzációs bejelentést szakmai körökben erős kétkedéssel fogadták.⁴ Annak az intelligens ágensnek, amelynek bármilyen esélye lehet a Turing-teszten, a következő képességekkel kell rendelkeznie:

- természetes nyelvi feldolgozás;
- tudásreprezentáció;
- automatikus következtetés;
- gépi tanulás.

¹ RUSSELL, Stuart J. – NORVIG, Peter: Artificial Intelligence: A modern approach. Pearson, Hoboken, New Jersey, 2021.

² RUSSELL, Stuart J. – NORVIG, Peter: Artificial Intelligence: A modern approach. p. vii.

³ BERNHARDT, Chris: Turing’s Vision: The Birth of Computer Science. The MIT Press, Cambridge, 2016. p. 157.

⁴ MASNICK, Mike: No, A ‘Supercomputer’ did NOT Pass the Turing Test for the First Time and Everyone Should Know Better. techdirt, 2014.06.09.
<https://www.techdirt.com/2014/06/09/no-supercomputer-did-not-pass-turing-test-first-time-everyone-should-know-better/>; letöltés: 2022.05.17.

Tekintettel a racionális – nem emberi – mesterséges intelligenciára, a fejlesztési folyamat során absztraktabb és egzaktabb megközelítések kerülnek alkalmazásra, mint például:

- matematika;
- statisztika;
- a formális logika számos ága, mint például:
 - Boole-logika;
 - elsőrendű logika;
 - deontikus logika;
 - fuzzy logika.

A jelenlegi tudományos közvélemény jobbra ezeket tekinti – többek között – az MI-kutatás és -fejlesztés tudományágainak. A robotika (a megtestesült mesterséges intelligencia) céljaira további tudományágakat és technológiákat fejlesztettek ki, mint például:

- számítógépes látás és arcfelismerés;
- beszéd felismerés;
- affektív számítástechnika az érzelmek kifejezésére (utánzás, emuláció).

Indeterminisztikus viselkedés: az alapvető kihívás

Egy hagyományos számítógépes programot úgy terveztek, hogy *determinisztikusan* működjön. A felhasználó azt várja el a szövegszerkesztőtől, az adatbáziskezelőtől, a levelező klientsztől és még egy sor általában használatos programtól, hogy „kreatív ötletek” nélkül, megfelelő sorrendben hajtsa végre az igényelt funkciókat. Azt – és csak azt – a műveletet végezze el, amire a felhasználó utasítja, és pontosan úgy, ahogy az utasítás szól. A program kiszámíthatatlan, előre nem látható viselkedése a kódban lévő hibára utal.

Egy mesterséges intelligencia alkalmazása azonban nem szokványos szoftver. Egyes MI-rendszerek persze determinisztikus működésű algoritmusokon alapulnak, de többnyire a mélytanuláson és a kapcsolódó technológiákon alapuló új rendszerek nem determinisztikus algoritmusokat valósítanak meg, és jó néhány tetszőleges elemet tartalmazó adathalmazt vonnak be a tanulásba. A valószínűségi működés e rendszereknek a velejárója. Ez azt jelenti, hogy *egy mesterséges intelligencia-rendszer nem determinisztikus módon működik*, és ez a döntő fontosságú tulajdonság magas biztonsági kockázatot jelent az MI-rendszer megvalósítása és használata során. Ezek a rendszerek a felhasználó által elvártnál kevésbé átlátható és megmagyarázható módon adhatnak válaszokat. Sok esetben még a mesterséges intelligencia szakértői és fejlesztői sem tudják pontosan megjósolni és teljes részletességgel megmagyarázni az MI viselkedését. Egy MI-rendszer valójában *fekete dobozként*⁵ értelmezhető; a felhasználó által adott bemeneti adatokra és utasításokra

⁵ TAN, Jonathan Ming En: Non-deterministic artificial intelligence systems and the future of the law on unilateral mistakes in Singapore. Singapore Academy of Law Journal, 34 SAclJ 91, 2022. p. 92. <https://journalonline.academyPublishing.org.sg/Journals/Singapore-Academy-of-Law-Journal/Current-Issue/ct/eFirstSALPDFJournalView/mid/494/ArticleId/1732/Citation/JournalsOnlinePDF>; letöltés: 2022.07.11.

a rendszer nem mindig nyilvánvaló és transzparens kimenetet produkál. A rendszer tényleges működését csak következtetési technikával lehet bizonyos szinten túl megközelíteni és megmagyarázni.

Ez a rendkívül kifinomult autonóm technológiák fejlesztésének járulékos – esetenként nem kívánatos – mellékhatása. A gépeink feletti ellenőrzés korlátozottsága teljesen új jelenség a technológiai civilizáció történetében, és még mindig nincsenek megfelelő válaszaink és irányelveink ennek a helyzetnek a kezelésére. A legfenyegetőbb és legriasztóbb forgatókönyv a mesterséges intelligencia és a robotika fejlődésében az *öntudatos mesterséges intelligencia* és esetleg az *emberfeletti képességekkel rendelkező rendszerek* megjelenése lenne. A tudományos viták és tanulmányok a sci-fi-műveken túl is figyelmeztetnek erre a lehetőségre.⁶ A biztonsági kockázat miatt az új technológia fejlesztése és a rájuk vonatkozó szabályozás kidolgozása különös figyelmet igényel.

Mára általánosnak tekinthető annak belátása, hogy elkerülhetetlen szükségszerűség a megfelelő ellenőrzések és biztonsági rendelkezések kidolgozása az MI-rendszerek fejlesztéséhez és használatához, politikai döntéshozó körökben is felerősödött az ezirányú motiváció. Ami a szabályozási kereteket illeti, bizonyára hamarosan új szabályozási doktrínákat kell kidolgoznunk a megszokott jogi eszköztáron túl. Az autonóm és intelligens rendszerek technológiai fejlődésével párhuzamosan egyre nyilvánvalóbbá válik annak szükségessége, hogy a mesterséges intelligenciát alkalmazó szolgáltatók és rendszerüzemeltetők szerepét és felelősségét megnyugtató és jövőálló módon szabályozzuk.⁷

A mesterséges intelligencia egyes alkalmazási területei

Az 1956-ban tartott Dartmouth Conference⁸ (Hannover, New Hampshire) emlékére az MI-kutatást megalapozó eseményként őrzi a tudományos világ. Az azóta eltelt évtizedek során a mesterséges intelligencia technológiai és módszerei sokat finomultak, fejlődtek, számos alkalmazási területen terjedtek el, és beszivárogtak a társadalmi, gazdasági és személyes tevékenységek számos szegmensébe. Néhány olyan ágazat, ahol a mesterséges intelligenciát széles körben alkalmazzák:⁹

- csillagászat és egyéb természettudományok;
- klimatológia;
- adat- és kiberbiztonság;
- e-kereskedelem;

⁶ TOTSCHNIG, Wolfhart: Fully Autonomous AI. Science and Engineering Ethics, Volume 26, 2020. pp. 2473–2485.

<https://doi.org/10.1007/s11948-020-00243-z>; letöltés: 2022.06.05.

⁷ CUSTERS, Bart – FOSCH-VILLARONGA, Eduard: Humanizing Machines: Introduction and Overview. In: CUSTERS, Bart – FOSCH-VILLARONGA, Eduard (szerk.): Law and Artificial Intelligence. Regulating AI and Applying AI in Legal Practice. T.M.C. Asser Press, The Hague, 2022. p. 10.

⁸ MCCARTHY, John – MINSKY, Marvin L. – ROCHESTER, Nathaniel – SHANNON, Claude E.: A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. August 31, 1955. <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>; letöltés: 2022.06.10.

⁹ Application of AI. javaTpoint, 2022.

<https://www.javatpoint.com/application-of-ai>; letöltés: 2022.08.11.

- oktatás;
- pénzügyek és bankszektor (tőzsdei irányítás és előrejelzés);
- szerencsejáték és szórakoztatás;
- egészségügy (diagnosztika);
- háztartási és személyes segítségnyújtás;
- gyártás;
- robotika;
- szociális média;
- közlekedés (navigáció, forgalomoptimalizálás, autonóm járművek).

Az említett szakterületek többsége szorosan kapcsolódik olyan tevékenységekhez, megnyilvánulásokhoz és érdekekhez, amelyek esetében a természetes személyek alapvető jogai és szabadságai vagy a vállalatok üzleti érdekei tekintetében közvetlen jogi és felelősségi kérdések merülhetnek fel a mesterséges intelligencia alkalmazásának okán.

A mesterséges intelligencia meghatározása az EU jogában

Az EU mesterséges intelligenciáról szóló törvénye (Artificial Intelligence Act – AIA)¹⁰ – amely technikailag egy készülő uniós rendelet tervezete – szintén kísérletet tesz a mesterséges intelligencia meghatározására a szabályozási keret céljaira. Az értelmező rendelkezések között található alábbi definícióról előljáróban megjegyzendő, hogy nem a „*mesterséges intelligenciát*” mint olyat, hanem a „*mesterségesintelligencia-rendszert*” határozza meg. Az absztrakt értelemben vett „mesterséges intelligencia” meghatározása sok etikai problémát és a kognitív tudományok területére tartozó szakkérdést vetne fel. A mesterségesintelligencia-rendszer ezzel szemben egyszerűbb, pragmatikusabb, technológiai oldalról jobban megragadható fogalom. A jogalkotó ezt az egyszerűbb utat választotta a 3. cikk. 1. pontjának szövegezésekor, amely szerint a „*mesterségesintelligencia-rendszer (MI-rendszer)*”: *olyan szoftver, amelyet az I. mellékletben felsorolt technikák és megközelítések közül egy vagy több alkalmazásával fejlesztettek, és amely az ember által meghatározott célkitűzések adott csoportja tekintetében olyan kimeneteket, például tartalmat, előrejelzéseket, ajánlásokat vagy döntéseket képes generálni, amelyek befolyásolják azt a környezetet, amellyel kölcsönhatásba lépnek.*

A definíció teljességéhez az említett I. melléklet megfelelő részének felhívása is szükséges, amely a fejlesztés során alkalmazott, mesterséges intelligenciára orientált technológiákat említi, ezek a következők:

¹⁰ Artificial Intelligence Act (AIA). Javaslat Az Európai Parlament és a Tanács Rendelete a mesterséges intelligenciára vonatkozó harmonizált jogszabályok (A Mesterséges Intelligenciáról szóló jogszabály) megállapításáról és egyes uniós jogalkotási aktusok módosításáról. Európai Bizottság, Brüsszel, 2021.04.21.
<https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52021PC0206&from=FR>;
letöltés: 2022.09.04.

- Gépi tanulási megközelítések, ideértve:
 - felügyelt,
 - felügyelet nélküli,
 - megerősítő tanulást, a módszerek széles skálája, többek között
 - mélytanulás alkalmazásával;
- Logikai és tudásalapú megközelítések, beleértve:
 - tudás megjelenítését;
 - induktív (logikai) programozást;
 - tudásbázisokat;
 - következtetőmotorokat;
 - szimbolikus érvelést és a szakértői rendszereket;
- Statisztikai megközelítések:
 - Bayes-féle becslés;
 - keresési és optimalizálási módszerek.

Ez az értelmezés a lehető legsemlegesebb, azaz a jogi koncepciók nem foglalnak állást abban a tudományos vitában, hogy a mesterséges intelligencia technikai kritériumait és minőségét az átlagos emberi készségekhez és képességekhez vagy egy absztrakt – matematikai és/vagy logikai – racionalitáshoz kell-e mérni.

A releváns technológiák felsorolása szintén a meghatározás érdemi részét képezi. Ez az értelmezés tükrözi a mesterséges intelligencia egyfajta tudományos meghatározását, taxonómiáját, s ugyanakkor segít megkerülni a mesterséges intelligencia kognitív természetével kapcsolatos kényes és megosztó kérdéseket.

A MESTERSÉGES INTELLIGENCIA RENDSZERTANA

A mesterséges intelligenciáról szólva nyilvánvaló, hogy többféle ilyen megoldást, rendszert különböztethetünk meg ezen a területen. Egyes rendszerek csak nagyon korlátozott képességekkel rendelkeznek, míg más rendszerek versenyre kelhetnek az emberi szakértők ismereteivel és viselkedésével. Az elmúlt évek során tiszteletre méltó eredményeket értek el az MI-rendszerek fejlesztői például az orvosi képalkotáson alapuló diagnosztikai eljárásokban, főként a rákdiagnosztikában, a terhességi tesztekben, az elektroenkefalográfiában stb. alkalmazták azokat.

A mesterséges intelligencia fejlesztésének jelenlegi állását figyelembe véve és némileg előre tekintve meghatározhatjuk a mesterséges intelligencia alkalmazásainak bizonyos kategóriáit. Ez a taxonómia a jogi gondolkodás szempontjából is szükséges, mivel a mesterséges intelligencia kiforrottsága és képességei meghatározó jelentőségűek lehetnek az ilyen rendszerekkel kapcsolatos felelősség jellegének, mértékének, esetleges korlátozottságának tekintetében. A mesterséges intelligencia kategorizálását a rendszer *funkcionalitása és képességei* alapján azonosíthatjuk.

Az MI rendszerezése funkcionalitás alapján¹¹

Reaktív mesterséges intelligencia

A reaktív mesterséges intelligencia úgy van programozva, hogy a kapott bemenet alapján determinisztikus kimenetet adjon. A reaktív motorokra épülő alkalmazások mindig minden alkalommal azonos módon reagálnak a bekövetkező helyzetekre, nem gyarapítják tapasztalatokkal a meglévő ismereteiket, azaz nem rendelkeznek a gépi tanulás képességével, valamint temporális, azaz a múltra és/vagy a jövőre vonatkozó kontextusokat kezelő, azokat figyelembe vevő képzetekkel sem.

Az ilyen típusú MI-szolgáltatások nem tudnak az eredetileg tervezett feladatokon túlterjeszkedve működni. Ez eleve korlátozottá teszi őket. Ami a működőképes példákat illeti, néhány jól ismert alkalmazást említhetünk:

- Deep Blue – az IBM sakkozó szuperszámítógépe, amelynek fejlesztett változata 1996-ban és 1997-ben két hatjátszmas tornát vívott Garry Kaszparovval, és a gép játékereje legalábbis összemérhető volt a világbajnokéval;
- e-mail-szerverekbe ágyazott SPAM-szűrő segédprogramok;
- Google/YouTube/Spotify/Netflix ajánlómotorja (*recommendation engine*).

Korlátozott memóriájú MI (Limited memory AI)

A korlátozott memóriával rendelkező mesterséges intelligencia tanul a múltból, és a cselekvések vagy adatok megfigyelésével empirikus tudást épít fel. Az ilyen típusú mesterséges intelligencia a múltbeli megfigyelési adatokat előre beprogramozott információkkal kombinálva használja fel előrejelzések készítéséhez és összetett osztályozási feladatok elvégzéséhez. Nyilvánvaló, hogy a gépi tanulási képességek a következő lépcsőfokot jelentik az előző pontban említett reaktív mesterséges intelligenciához képest. Ma már ezek a rendszerek is nagyon elterjedtek. Megjegyzendő, hogy az autonóm járműveket irányító MI-rendszerek szintén az ilyen típusú alkalmazások egyik sajátos formáját jelentik.

A korlátozott memóriájú mesterséges intelligencia, ahogy a neve is mutatja, még mindig korlátozott. Az információ, amellyel az autonóm járművek dolgoznak efemer jellegű, és nem tárolódik az autó hosszú távú memóriájában.

Tudatelméleti mesterséges intelligencia (Theory of mind AI)

Az ebbe a kategóriába tartozó mesterségesintelligencia-rendszerekre még nem találunk ipari példákat. Jobbára csak néhány tudományos és technológiai kísérlet említhető, amelyekben a döntéshozatali képességek néhány kezdetleges eleme megegyezik az emberével – vagy zavarba ejtően hasonlít hozzá. Az ilyen kognitív mesterséges intelligenciával rendelkező gépek képesek lesznek megérteni és megjegyezni az érzelmeket, majd az emberekkel való interakció során ezen érzelmek alapján beállítani, finomhangolni saját viselkedésüket.

¹¹ MARR, Bernard: What are the Four Types of AI? Bernard Marr, 2021.
<https://bernardmarr.com/what-are-the-four-types-of-ai/>; letöltés: 2022.09.01.

Egy intelligens beszélgetés egy érzelmileg intelligens robottal, amely úgy néz ki és úgy szólal meg, mint egy valódi ember, ezekkel a gépekkel már megvalósítható lesz. A tudatosságon alapuló mesterséges intelligencia megvalósítása előtt azonban még sok akadály áll, mivel az emberi kommunikációban a gyorsan változó érzelmek alapján történő viselkedésváltoztatás folyamata nagyon rugalmas. Ezt pedig nagyon nehéz imitálni, amikor érzelmileg egyre intelligensebb gépet próbálunk létrehozni.

Néhány humanoid robot – mint például a hongkongi Hanson Robotics által kifejlesztett Sophia – képes demonstrálni az emberi felhasználókkal való társas interakciók bizonyos képességeit. „Ő” képes felismerni az arcokat és a saját arckifejezésével reagálni az interakciókra.

Öntudatos mesterséges intelligencia (Self-aware AI)

Ez a kategória jelenleg csak a sci-fi világában létezik, és nem lehet tudni, legfeljebb becsülgetni, hogy mikor jelenik meg a mesterséges intelligenciának ez a rendkívül fejlett formája. Jelenleg nem rendelkezünk az ehhez szükséges hardverrel, és nem ismerjük azokat az algoritmusokat sem, amelyek egy ilyen gépet működésre bírnanak. Ez a mesterséges intelligencia egy olyan gép, amely öntudatos és saját érzelmekkel rendelkezik, és nem csak arra képes, hogy – többé-kevésbé – helyesen reagáljon a vele kapcsolatban álló emberek cselekedeteire és érzelmeire.

Ez a fajta mesterséges intelligencia nemcsak öntudatos lesz, hanem vágyakkal, szükségletekkel és érzelmekkel is rendelkezik majd.

Emberfeletti mesterséges intelligencia (Superhuman AI)

Ez a kategória még a sci-fi-irodalom és -film mértéktartóbb kiadásain is túlmutat, de sajátos módon már ma is tudományos viták tárgya. Csak találgatások és sejtések lehetnek erről a lényről, és egészen homályos előérzetek. Bizonyos állítások azt jósolják, hogy az emberfeletti mesterséges intelligencia kifejlesztése fizikailag lehetséges. Joseph Carlsmith egy vaskos tanulmányt szentelt ennek a forgatókönyvnek, és 2070-re jósolja a valószínűségét.¹²

Ugyanebben a próféciában a szerző arra a következtetésre jut, hogy egy ilyen rendszer létrejötte akaratlanul is az összes ember végleges ellehetetlenítéséhez, feleslegessé válásához vezetne, ami az emberiségre nézve egzisztenciális katasztrófát okozna.¹³

A mesterséges intelligencia tipológiája a képességek alapján¹⁴

A mesterséges intelligencia másik megszokott és széles körben alkalmazott taxonómiája e rendszerek képességein alapul. E megközelítés szerint a mesterséges intelligenciának az alábbi típusai különböztethetők meg.

¹² CARLSMITH, Joseph: Is power-seeking AI an existential risk? Open Philanthropy, April 2021. p. 13. <https://docs.google.com/document/d/1sma11lagHHcrhoi6ohdq3TYIZv0eNWWZMPEy8C8byYg/edit#heading=h.pwdbumje5w8r>; letöltés: 2022.08.12.

¹³ CARLSMITH, Joseph: Is power-seeking AI an existential risk? p. 64.

¹⁴ Types of Artificial Intelligence. javaTpoint, 2022. <https://www.javatpoint.com/types-of-artificial-intelligence>; letöltés: 2022.08.11.

Gyenge vagy szűk mesterséges intelligencia (Weak AI or Narrow AI)

A gyenge (vagy szűk) mesterséges intelligencia az MI olyan típusa, amely képes egy adott feladat intelligens végrehajtására. Jelenleg ezek a legelterjedtebb és elérhető MI által támogatott vagy működtetett rendszerek. A gyenge MI nem képes a saját területén vagy korlátain túl teljesíteni, mivel csak egy adott feladatra tervezték és képezték ki. Az adott tartományon túl a gyenge MI működése megbízhatatlan és kiszámíthatatlan.

Tekintsünk néhány olyan példát, amelyet már széles körben használnak:

- az Apple Siri egy gyenge mesterséges intelligencia, amely korlátozott, előre meghatározott feladatkörrel működik;
- az IBM Watson online *soft-computing* eszközök szintén a Narrow AI-t képviselik, Watson képességei a következők:
 - szakértői rendszermegközelítés;
 - gépi tanulás;
 - természetesnyelv-feldolgozás.
- Egyéb gyenge MI-alkalmazások:
 - sakk – és más stratégiai – játékok;
 - vásárlási ajánlómotorok e-kereskedelmi oldalakon;
 - autonóm járművek;
 - beszédfelismerő berendezések;
 - képfelismerő rendszerek.

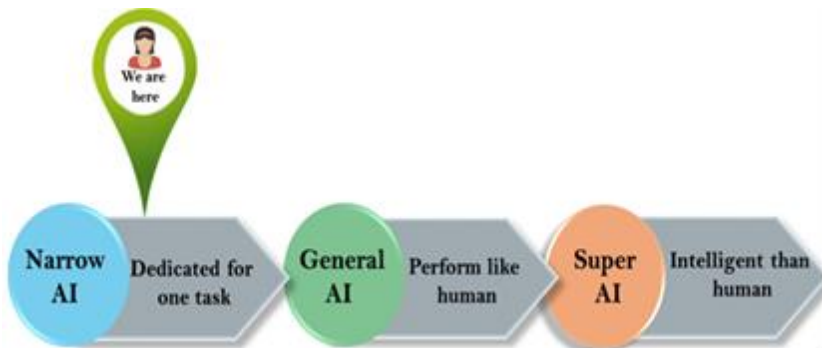
Általános mesterséges intelligencia, ÁMI (Artificial General Intelligence)

Az általános mesterséges intelligencia az intelligens ágensek egy magasan fejlett típusa, amely képes bármilyen szellemi feladatot emberhez hasonló teljesítménnyel megoldani és végrehajtani. Egy ÁMI-rendszer intelligensebb lenne, és önmagában úgy gondolkodna, mint egy ember. Jelenleg nem létezik ilyen rendszer, de ez az MI-kutatás és -fejlesztés egyik elsődleges célja. Ennek a kutatási erőfeszítésnek az időtartama megjósolhatatlan. Az MI-szakértők, tudásmérnökök többnyire egyetértenek abban, hogy az ÁMI-nek rendelkeznie kell a következő képességekkel:

- tudásreprézntáció;
- gondolkodás;
- stratégiafejlesztés;
- feltételes valószínűségi döntéshozatal (Bayes-féle problémák megoldása);
- tervezés;
- tanulás (gépi tanulás);
- természetes nyelvi kommunikáció;
- mindezen készségek integrálása.

Erős vagy szuper MI

Az erős vagy szuper MI egyelőre inkább hipotetikus fogalom. A gépi intelligencia olyan szintjére utal, amely meghaladhatja az emberi készségeket és kognitív képességeket. Ez az ÁMI egyik eredménye lenne. A pesszimista – vagy netán realista – előrejelzések szerint ennek megvalósulása katasztrofális jövőt jelentene az emberiség számára.



1. ábra. Az MI három típusa¹⁵

A MESTERSÉGES INTELLIGENCIÁVAL KAPCSOLATOS JOGI KOCKÁZATOK

Hosszan lehetne sorolni a mesterséges intelligenciával kapcsolatos pesszimista várakozásokat, forgatókönyveket. Az emberek jó részében azzal kelt szorongást a mesterséges intelligencia, hogy az emberi mivoltunk legbensőbb lényegének érzett attribútummal, a gondolkodásra, a racionális – tudatosnak tartott – döntésre és cselekvésre irányuló képességgel kel versenyre. Ezt a nyugtalanságot csak fokozza a tapasztalat, hogy az eddig megvalósult MI-fejlesztések a maguk – egyelőre szűk és specializált – működési körében gyakran az embert felülmúló teljesítményt mutatnak. Egy nagy teljesítményű szkennel és OCR-rendszer gyorsabban digitalizál nyomtatott szöveget, mint amilyen sebességgel az emberi felhasználó olvasni képes.

A DeepBlue – már megint ezt az esetet említhetjük – sokkal jobban sakkozik, mint egy átlagember; játékeréjével csak a legnagyobb bajnokok dacolhatnak. De a jóval a szupergép alatti szinten álló Shredder Chess sakkprogram is keményen próbára tesz egy átlagosan jó emberi játékost. Igazán riasztó tapasztalatok.

A mindennapokból jól ismert számítógépes fordító szolgáltatások – bár még el-el követnek komikus nyelvtani hibákat is – a poliglottokat megszegyenítő képességekkel rendelkeznek. Ráadásul egészen nyilvánvaló, hogy amit a mesterséges intelligencia egyszer megtanult, azt nem felejt el, és a már megszerzett képességeket sem kell újra megtanulnia, tehát a tudása folyamatosan gyarapszik, képességei egyre mélyülnek, és

¹⁵ What Are The 3 Types Of AI? Deccan Herald, 2020.06.24.
<https://www.deccanherald.com/brandspot/pr-spot/what-are-the-3-types-of-ai-853275.html>;
letöltés: 2022.08.10.

pillanatnyilag nem tudunk olyan határról, amelyen túl ne növekedhetne. Ezzel az adottsággal az élet végességébe zárt ember nem lehet képes lépést tartani. Hiszen minden kisgyerek a teljes tanulási folyamatot előlről kezdi, és élete során a felhalmozott egyetemes tudásnak csak ijesztően kis hányadát sajátíthatja el.

Egy ilyen – részben futurisztikus – áttekintés után most láthatjuk, hogy milyen kihívásokkal és kockázatokkal kell számolnunk ma a mesterséges intelligencia alkalmazásával kapcsolatban. Melyek azok a reális kockázatok, károk és sérülések, amelyeket az MI-rendszerek az egyének, az egyének csoportjai és a társadalom egésze számára okozhatnak? Milyen típusú mesterségesintelligencia-rendszerekkel van gyakorlati tapasztalatunk?

A mesterséges intelligenciának csak a következő kategóriái tekinthetők létező rendszereknek és szolgáltatásoknak:

- reaktív mesterséges intelligencia;
- korlátozott memóriájú mesterséges intelligencia;
- gyenge/szűk mesterséges intelligencia.

A többi említett kategória egyelőre a fantázia világába tartozik, és tisztán kell látnunk, hogy a sci-fi nem a jog valódi működési területe. A jogi gondolkodásnak következőképpen a gyakorlatban ténylegesen létező és működő MI-rendszerek által támasztott kihívásokra kell koncentrálnia.

Ezek az intelligens ágensek alsóbb osztályába tartoznak, de ezeket is érdemes a jogi kockázatok forrásaként kezelni. Anélkül, hogy a mesterséges intelligenciával működő rendszerek által veszélyeztetett jogi érdekek átfogó felsorolására törekednénk, könnyen azonosíthatunk néhány alapvető kategóriát.

Jogi megfontolást igénylő mesterséges intelligenciával kapcsolatos károk és kockázatok az alábbi fő területeken merülhetnek fel:

- *személyiségi jogok* – különösen a személyes adatok sérelme révén;
- *szellemi tulajdonnal* kapcsolatos jogok, különösen szerzői jogi és védjegyjogi területen;
- *vagyoni jogok*, ha a mesterséges intelligencia működése harmadik személy számára anyagi kárt okoz (pl. autonóm jármű által okozott baleset; MI felügyelete alatt működő rendszer hibájából fakadó vagyoni kár).

A kockázatok eredete és természete

A gyakorlatban működő mesterségesintelligencia-rendszereket vizsgálva számos olyan tényezőre akadhatunk, amelyek növelik a kockázatokat és jogi szabályozást, rendezést kívánnak.

Gyors piaci növekedés

Az első és talán legfeltűnőbb tényező az MI-alkalmazások, -szolgáltatások és -termékek piacának viharos gyorsasággal történő bővülése. Az utóbbi néhány évben ez a folyamat – kutatási és fejlesztési aprómunkával eltelt sok évtized után – hirtelen

felgyorsult és igen látványossá vált. Tekintélyes elemzőintézetek között izzott fel a „számháború” az MI-piac kapitalizációjának mértékéről és növekedési kilátásairól. Tüzetes ellenőrző kutatások nélkül is bizvást megállapíthatjuk, hogy imponáló adatokról van szó. Csak egy példa a sok közül a *MarketsandMarkets*¹⁶ elemzése, amely 2022-re 86,9 milliárd dolláros piacméretet regisztrált, 2027-re pedig ennek az ötszörösét, 407 milliárd dollárt prognosztizál, ami 36,2%-os átlagos évenkénti növekedési rátát jelent. A mesébe illő növekedést – nem nagy meglepetés – több tényező együttes hatásának tulajdonítják az elemzők:

- növekvő igény az üzleti életben az intelligens rendszerek iránt számos területen:
 - döntéstámogatás;
 - valós idejű vásárlási tanácsadás;
 - interaktív játékok.
- a kibertámadások, biztonsági incidensek számának emelkedése által megnövelt igény az intelligens biztonsági rendszerek iránt;
- történeti adathalmazok megnyitása, amelyek hosszú távú trendszámításokat tesznek lehetővé.

Egy ilyen viharos sebességű növekedés a történeti tapasztalatok szerint mindenképpen gazdasági kockázatok sokaságával jár. Ha azonban egy olyan érzékeny és különleges közegben következik be, mint az MI-szolgáltatások ökoszisztémája, akkor a kockázatok fokozódnak. A gyorsan változó gazdasági és szabályozói környezetben még hiányoznak a kellő tapasztalatok, kevés a megfelelően felkészült szakember, az ellenőrző szervezetek, hatóságok esetlenül mozognak az ismeretlen terepen, tehát gyengül a munkájuk hatékonysága, és a közvélemény sem kellően tájékozott a józan önvédelemhez.

Az mindenképpen előre látható, hogy a trendek szerint viszonylag rövid időn, néhány éven belül a gazdasági élet szereplőinek és az állami hatóságoknak jelentős része alkalmazni fog valamilyen – vélhetően egyszerre többféle –, mesterséges intelligenciára épülő rendszert. Ezek a rendszerek pedig közvetve vagy közvetlenül a társadalommal, annak alrendszerével és egyes természetes személy tagjaival kapcsolatos feladatokat látnak el. Ez tehát egy új gazdasági, technológiai és jogi ökoszisztémának a felépülésével jár, aminek a körvonalai éppen csak kibontakozóban vannak. Bizvást állíthatjuk, hogy az MI-alapú rendszerek hamarosan éppen olyan szervesen és mélyen átjárják a globalizált piaci és politikai rendszert, amilyen rohamos gyorsasággal a digitális adatkommunikáció, az internet alapszolgáltatásai az ezredfordulón meghódították a világot.

¹⁶ Artificial Intelligence Market by Offering (Hardware, Software, Services), Technology (Machine Learning, Natural Language Processing), Deployment Mode, Organization Size, Business Function (Law, Security), Vertical and Region – Global forecast to 2027. MarketResearch, August 2022. <https://www.marketresearch.com/MarketsandMarkets-v3719/Artificial-Intelligence-Offering-Hardware-Software-32075223/>; letöltés: 2022.07.21.

Átláthatósági problémák

Az MI-rendszerek nyilvánvaló inherens kockázata a nem, vagy csak nehezen átlátható működés. Bár a mesterséges intelligenciát működtető algoritmusok ismert, megismerhető tudományos tételeken alapulnak, a rendszerek betanítása olyan adathalmazok felhasználásával történik, amelyekben sok az esetlegesség. A jelenlegi BigData-korszakban, amikor a döntéshozatalhoz hatalmas mennyiségű adat vált elérhetővé, azt tapasztalhatjuk, hogy az MI-algoritmusok végső soron kockázatokat jelentenek a társadalomra és az egyénekre nézve. A gyors – és minden bizonnyal tovább gyorsuló – technológiai fejlődéssel azonos ütemben kibontakozó tudományos párbeszédben a mesterséges intelligencia aktuális vitatémává vált. Ebben a diskurzusban az *átláthatóság garantálásához*, illetve az *üzleti titkok védelméhez* fűződő versengő érdekek nyilvánulnak meg.

Egyfelől nyilvánvaló társadalmi érdek szól a transzparens módon működő, tisztességes és megbízható MI-rendszerek iránti közbizalom megteremtése mellett. Ugyanakkor ennek a törekvésnek ellentmond a fejlesztő szintén méltányolható érdeke és igénye a művével kapcsolatos speciális szakmai tudás bizalmosságának védelme iránt. Az üzleti titkok megőrzésének célja a kereskedelmi érdekek védelme és az egyének, a vállalatok és végső soron az Unió gazdasági fejlődésének ösztönzése.

Az átláthatatlanság, az algoritmusok és a rendszer egészének üzleti titokként történő kezelése olyan ártalmak és kockázatok lappangó jelenlétének ágyaz meg, mint a diszkrimináció, a tisztességtelenség, a magánszféra bizalmosságának sérelme. Ezt nyilván ellensúlyozni kell. Az átláthatóság az MI-vel kapcsolatos fő aggályok, a tisztességtelenség, a megkülönböztetés és az átláthatatlanság elkerülésére törekszik. Tágabb értelemben az *algoritmikus átláthatóság* követelményének érvényesítése magában foglalná a jogot a tájékoztatáshoz és az adatok tisztességes feldolgozásához, valamint a közérdek védelmét.

A jog ezt olyan eszközökkel, szabályozási doktrínákkal, alapelvekkel kísérli meg, mint a sokat vitatott, hányatott sorsú és végül a GDPR végleges szövegében is csak felemás módon megjelenő „*magyarázathoz való jog*”, amely szerint az adatalanyok joga lenne magyarázatot igényelni az adatkezelőtől az automatizált egyedi döntéssel. Ez a megoldás sem nyújt megnyugtatóan szilárd jogi garanciákat, az átláthatatlan működés a személyes adatokat érintő adatfeldolgozásokon túlmenően is jellemzi a mesterséges intelligenciát.

Előítéletes MI

A mesterségesintelligencia-rendszerek nem izoláltan, hanem egy társadalmi környezetbe ágyazva működnek, és döntéseik eredményei is ebben a közegben realizálódnak. Mára közismertnek tekinthető az a tapasztalat, hogy egyes rendszerek a társadalom különböző csoportjaihoz különbözően, előítéletesen, akár diszkriminatív módon viszonyulnak. Röviden ezt nevezhetjük a mesterséges intelligencia elfogultságának, részrehajlásának is. Legfajóbban azokban az esetekben nyilvánul meg, amikor emberek, embercsoportok életére, sorsára nézve fontos döntések múlnak a mesterséges intelligencia működésén, választásain.

Ennek a nyilvánvalóan káros jelenségnek az okait kutatva számos tényező ismerhető fel. Az Amerikai Egyesült Államok Szabványügyi Intézete (National Institute of Standards and Technology) a 2022-ben publikált jelentésében¹⁷ a mesterséges intelligencia elfogultságának alábbi három fő kategóriáját azonosítja:

- rendszerszintű előítéletek;
- statisztikai és számítási torzítások;
- emberi előítéletek.

A rendszerszintű előítéletek bizonyos intézmények eljárásaiból és gyakorlataiból erednek, amelyek olyan módon működnek, hogy bizonyos társadalmi csoportok előnyben részesülnek, mások pedig hátrányos helyzetbe kerülnek vagy leértékelődnek. Ennek legkirívóbb esetei az intézményesült rasszizmus és szexizmus. Figyelmeztető és riasztó jel, hogy ezek az oly sok tragédiáért felelős előítéletek olyan erősen gyökereznek a társadalmakban, hogy még a mesterséges intelligencia – sterilnek tűnő – világa sem tud mentesülni tőlük.

A statisztikai és számítási torzítások olyan hibákból erednek, amelyek akkor keletkeznek, amikor a betanításhoz használt adatminta nem volt reprezentatív a populációra nézve. A hiba származhat az adatok heterogenitásából, az összetett adatok túlzottan leegyszerűsített matematikai ábrázolásban történő megjelenítéséből, téves adatok alkalmazásából és algoritmikus torzításokból.

Az emberi előítéletek az emberi gondolkodás öröklött, gyakran kulturális meghatározottságú hibáit tükrözik. Ezek lehetnek az emberi döntéshozatalt jellemző heurisztikus elvek, vagy egyszerűbb ítélkezési műveletek esetén a kimenetel intuitív előrejelzése.¹⁸ Ezek az előítéletek gyakran implicitek, és általában ahhoz kapcsolódnak, hogy egy egyén vagy csoport hogyan érzékeli az információt (pl. az automatizált mesterséges intelligencia kimenetét), hogy döntést hozzon vagy kitöltse a hiányzó vagy ismeretlen információt.

A lehetséges következmények

Érzékeny és veszélyes technológia kiterjedt alkalmazásának nem is lehet más következménye, mint az, hogy időről időre súlyos hibák lépnek fel, amelyek károkat okoznak. Természetesen még ezekkel a károkkal és kockázatokkal együtt is azt tapasztaljuk, hogy a mesterséges intelligencia használata és további fejlesztése jelentős gazdasági, kulturális értéket hoz létre, tehát többet használ, mint amennyit árt; nyilván ebben áll a vonzereje. Fordított esetben nem volna érdemes foglalkozni vele, és tartózkodni kellene a használatától.

¹⁷ SCHWARTZ, Reva – VASSILEV, Apostol – GREENE, Kristen K. – PERINE, Lori – BURT, Andrew – HALL, Patrick: Towards a Standard for Identifying and Managing Bias in Artificial Intelligence. NIST Special Publication 1270. NIST, 2022.03.15. pp. 6–7.
https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=934464; letöltés: 2022.06.25.

¹⁸ SCHWARTZ, Reva – VASSILEV, Apostol – GREENE, Kristen K. – PERINE, Lori – BURT, Andrew – HALL, Patrick: Towards a Standard for Identifying and Managing Bias in Artificial Intelligence. p. 8.

Az összességében – egyelőre – pozitív mérleg mellett a társadalomnak fel kell készülnie azokra a váratlanul bekövetkező károokra, incidensekre, üzemzavarokra, amelyek a mesterséges intelligencia alkalmazásával függenek össze. Ezek következménye a kiterjedt alkalmazás miatt tömegesen fellépő, sokakat elérő negatív hatású esemény lehet:

- jelentős vagyoni kár bekövetkezése;
- tömeges személyiségi jogsérelem, például a Facebook–Cambridge Analytica-botrány, amelynek mintegy 87 millió Facebook-felhasználó¹⁹ volt a sértettje;
- politikai következmények is előállhatnak, hiszen a mesterséges intelligencia profilképző erejével felruházott média tömegek véleményének hatékony alakítója, így politikai események – mondjuk választások – eredményének eldöntője is lehet.

Veszélyeztetett személyiségi jogok

Profilalkotás és web-scraping

Az adatvédelmi jog az európai jogi kultúra jelentős újítása, amely az informatika területén bekövetkezett fejlődés következményeként néhány évtized alatt a személyiségi jogok védelmének élvonalába került. Nemzeti adatvédelmi hatóságok, jogvédő civil szervezetek, aktivisták küzdenek a modern állam és az interneten tevékenykedő adatkereskedő cégek, közösségimédia-szolgáltatók információs túlhatalma ellen.

A digitális technológia – a mesterséges intelligencia képességeivel megerősítve – kifinomult eszközöket és módszereket biztosít az adatkezelők számára a társadalom és a magánszemélyek megfigyelésére és a személyiségüket részletesen leíró, reprezentáló személyiségprofil megalkotására. Az emberiség történelmében soha egyetlen államnak vagy más szervezetnek sem voltak olyan hatékony eszközei a társadalom ellenőrzésére és manipulálására, mint amilyeneket ma a mindennapi használatban láthatunk.

A közösségimédia-szolgáltatók, a webáruházak és a kormányzati szervek profilalkotási képességei többnyire MI-algoritmusokon alapulnak. A médiában riasztó tudósításokat láthatunk, olvashatunk tömeges adatvédelmi incidensekről. A Facebook–Cambridge Analytica-adatbotrány jól szemlélteti a masszív algoritmikus profilalkotás társadalmi és politikai kockázatát. A GDPR és más kapcsolódó jogszabályok célja pedig az, hogy biztonságos, méltányos és törvényes kereteket biztosítsanak ezekkel a visszaélésekkel és bűncselekményekkel szemben.

¹⁹ KOZŁOWSKA, Hanna: The Cambridge Analytica scandal affected nearly 40 million more people than we thought. Quartz, 2018.04.04.
<https://qz.com/1245049/the-cambridge-analytica-scandal-affected-87-million-people-facebook-says/>
letöltés: 2022.05.30.

Ezek az incidenseken túl azonban fennáll annak a kockázata is, hogy az MI-algoritmusok által feldolgozott és profilozott²⁰ adatok téves következtetésekhez és ezáltal megalapozatlan döntésekhez vezethetnek. A jognak választ kell adnia arra, hogy ki és hogyan felel az ilyen hibák, illetve visszaélések által okozott károkért.

Az intelligens szoftverek által végzett *webes adatgyűjtés*, közismert nevén a *web-scraping* egyre szélesebb körben elterjedt gyakorlat, amely szintén különleges adatvédelmi kockázatot jelent.

Deepfake videók – mint személyiségi jogsértés

A *deepfake* kifejezés és jelenség – az a kárhozatos technikai trükk, hogy egy filmfelvételen az elkövető a valódi szereplő helyére – jellemzően az arc helyére – más személynek a vonásait illeszti be általában lejáratási szándékkal – rövid idő alatt futott be jelentős karriert. Bár a jelenség előképei már a mozgóképes kultúra korai szakaszáig visszavezethetők, technikai okok miatt ez csak korlátozottan, igen sok emberi munkával és jellemzően rossz minőségben volt kivitelezhető a fotokémiai alapú filmkészítési technológiák korában.

A digitális fényképezés és mozgóképrögzítés, valamint az MI egymásra találása azonban hatalmas új távlatokat nyitott a rosszindulatú, befeketítő felvételhamisítás előtt. Ma lényegében bárki vizionálhatja magát vagy ismerőseit mondjuk egy pornográf felvételen, vagy egyéb olyan helyzetben, ami becsületének, jó hírnevének súlyos csorbítására alkalmas. Az elkövetőnek még csak jelentős technikai tudással sem kell rendelkeznie, hiszen a mozgókép-manipulációhoz szükséges szoftverek könnyen elérhető termékként beszerezhetők az informatikai piacról. Ezekben az eszközökben fejlett MI-alapú eszközök végzik képek meggyőző retusálását. A kontúrok, a mozdulatok, a mimika természetessége, a hangsávnak a képhez történő igazítása, szinkronizálása ma már olyan jó minőségű hamisítványok készítését teszi lehetővé, ami a felületes szemlélőt könnyen megtéveszti, azaz meggyőzi a látvány valóságáról. A technológia gyors fejlődése révén a minőség nyilván tovább fog javulni, ami nem jó hír a hamisításnak kitétt személyek – vagyis gyakorlatilag bárki – számára.

A *deepfake* első sértettjei jellemzően ismert emberek, közéleti, művészeti celebritások voltak, köztük sok filmszínész, akik tipikusam pornográf jelenetek ártatlanul hírbe hozott szereplőiként tűntek fel. Az ilyen jellegű felvételeknek az első tömeges közzétételi platformja a *Reddit* közösségi hálózat volt, ahol 2017-ben született meg maga a *deepfake* elnevezés is, egy *Reddit*-felhasználó accountjából köznévvé válva.²¹

²⁰ CHAN, Rosalie: The Cambridge Analytica whistleblower explains how the firm used Facebook data to sway elections. Insider, 2019.10.05.

<https://www.businessinsider.com/cambridge-analytica-whistleblower-christopher-wylie-facebook-data-2019-10>; letöltés: 2022.07.04.

²¹ COLE, Samantha: We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now. A user-friendly application has resulted in an explosion of convincing face-swap porn. VICE, 2018.01.24.

<https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley>; letöltés: 2022.08.12.

A politikusok sem úszták meg a videómanipulációt. Az egyik első híresség Barack Obama korábbi amerikai elnök volt, aki egy 2017-ben közzétett módosított videófelvételen olyan szöveg szavait artikulálja igen meggyőző szájmozgással, amelyek egy másik felvétel hangsávján szerepelnek. Obamának annyiban szerencséje volt, hogy ez a módosított felvétel egy tudományos kutatási projekt, a *Synthesizing Obama* keretében készült, és nem ártani akartak vele az egykori elnöknek, hanem személyén keresztül demonstrálni a technológia képességeit.²²

A *deepfake mozgalom* azzal kezdődött, hogy a pornográf felvételekbe hírességek arcát illesztették be, a gyakorlat azonban hamarosan áttért átlagemberek, ismerősök, barátok, volt osztálytársak képeinek használatára. A mára betiltott *Deepfakes* subreddit és a *Discord* chatroom volt a melegágya annak, hogy a felhasználók az ilyen tolakodó vagy becsületsértő pornóvideók készítésére vonatkozó tippeket cseréljenek.

A technológia rétegei itt látványosan rakódnak egymásra. Vannak olyan nyílt forráskódú *web-scrapers*²³ eszközök, amelyek segítségével a felhasználók könnyen letölthetik egy személy közösségi média fiókjának összes képét, hogy létrehozzák azt az arckészletet, amit *deepfake* videó készítéséhez fel tudnak használni. A meggyőző *deepfake* elkészítéséhez azonban a készítőnek olyan testet kell találnia, amely könnyen illeszthető az áldozat arcához.²⁴ Az ideális test megtalálása is kvázi automatizáltá vált. Az arcfelismerő mesterséges intelligenciát alkalmazó böngészőalapú keresőmotorok az elkészítendő hamis videóban megjelenítendő sértett személy fényképe alapján keresik meg azt a felvételt, amelyben a célszemélyhez leginkább hasonló pornósztár szerepel. A *deepfake* videó ezekből a komponensekből szinte néhány kattintással elkészíthető – a személyiségi jogok súlyos sérelmére.

Személyiségi jogok védelme

A személyiségi jogok védelme, amely legalább 200 éves múltra tekinthet vissza, a modern kor jelentős jogi és kulturális fejleménye. Az évszázadok során azonban a jog a személyiség bonyolult és finom szövetének csak egyes pontjait, a személyiség néhány, szinte öletszerűen kiválasztott objektívációját részesítette explicit jogi oltalomban.

A digitális adatfeldolgozás és kommunikáció tömeges elérhetővé válása ebben a tekintetben is új korszakot hozott. Világossá vált, hogy a személyiség igen sérülékennyé vált, magánszférája a róla szóló adatok sokasága révén igen alaposan feltérképezhető, kiismerhető, profilozható, a személyiségi jogok védelmének klasszikus eszköztárára pedig már elégtelen a kiépülő információs túlhatalom kordában tartására, a személyiség magánszférájának adekvát biztosítására. Ennek az

²² SUWAJANAKORN, Supasom – SEITZ, Steven M. – KEMELMACHER-SHLIZERMAN, Ira: Synthesizing Obama: learning lip sync from audio. *ACM Transactions on Graphics*, Volume 36, Issue 4, 20 July 2017. pp. 1–13. <https://dl.acm.org/doi/10.1145/3072959.3073640>; letöltés: 2022.07.24.

²³ DownAlbum. Chrome Stores. <https://chromestores.com/extension/downalbum-extension/>; letöltés: 2022.08.25. Instagram Scraper. APIFY. <https://apify.com/apify/instagram-scraper#need-something-different-but-still-customizable>; letöltés: 2022.08.23.

²⁴ COLE, SAMANTHA: People Are Using AI to Create Fake Porn of Their Friends and Classmates. *VICE*, 2018.01.26. <https://www.vice.com/en/article/ev5eba/ai-fake-porn-of-friends-deepfakes>; letöltés: 2022.08.12.

új kihívásnak a kezelésére született meg az 1970-es évek elején a modern adatvédelmi jog intézményrendszere.

A technológia fejlődésével az adatvédelmi jog igyekszik lépést tartani, de a jogrendszer természetéből fakadóan állandóan csak követő pozícióban tud lenni. Előbb megszületik a technológia, amiből gazdasági, kulturális, társadalmi következmények fakadnak, s ezek a következmények – amelyek esetenként kifejezetten károsak is tudnak lenni – igényelnek megfelelő jogi kereteket.

A magyarázathoz való jog

A digitális technológia gyors fejlődésére érzékenyen figyelő jogpolitikai és jogalkotás néhány éve észlelte a mesterséges intelligenciával kapcsolatos kihívások gyarapodását, s első ösztönös reakciójaként az adatvédelmi jog intézményrendszerének bővítésével válaszolt. Kifejezetten az MI-rendszerek támasztotta adatvédelmi problémák kezelésének sajátos eszközeként merült fel a *magyarázathoz való jog* intézményének kidolgozása. Lényege szerint azt a lehetőséget biztosítaná az adatalannak, hogy az adatkezelőtől – aki alkalmasint mesterséges intelligenciát használ az adatok feldolgozása során – részletes magyarázatot igényelhessen az adatfeldolgozásról.

Ez a gondolat jól illeszkedett a GDPR akkor zajló kodifikációs folyamatába. Figyelemre méltó, hogy a GDPR nem nevezi nevén a mesterséges intelligenciát. Ehelyett egy semlegesebb hangzású és szélesebb tartalmú kifejezéssel „*kizárólag automatizált adatkezelésen alapuló döntésnek*” nevezi a mesterséges intelligencia működését, és általában az emberi beavatkozás nélküli adatfeldolgozást.

Az ilyen döntésekkel kapcsolatban az adatalannak igénye lehet arra, hogy magyarázatot kérjen és kapjon a döntés körülményeiről, a kiértékelési algoritmusról, a felhasznált adatokról. A listát kissé parttalanul folytathatnánk, ugyanis a koncepcióban a kodifikáció korai szakaszában még benne lévő *magyarázathoz való jog* végül kimaradt a végső változatból. Így nincs olyan kötelező erejű rendelkezés, ami ennek a jogosultságnak a pontos tartalmát meghatározná.

Az a furcsa helyzet állott elő, hogy a magyarázathoz való jog a GDPR-ban csak az egyébként kötelező erővel nem bíró Preambulum 71. bekezdésében szerepel ezzel az explicit megfogalmazással:

„Az ilyen adatkezelés mindazonáltal csakis megfelelő garanciák mellett végezhető, amelybe beletartozik az érintett külön tájékoztatása és az ahhoz való joga, hogy emberi beavatkozást kérjen és kapjon, különösen hogy kifejtse álláspontját, hogy magyarázatot kapjon az ilyen értékelés alapján hozott döntésről és hogy megtámadja a döntést.”

A GDPR megalkotása során több szövegváltozat készült, s ezek némelyikében a kötelező rendelkezések között szerepelt a magyarázathoz való jog is, ám a végső szövegből a jogalkotási alkudozás során ez a rendelkezés kimaradt,²⁵ és csak egy enyhébb kötelezettségeket előíró rendelkezés maradt meg.

²⁵ WACHTER, Sandra – MITTELSTADT, Brent – FLORIDI, Luciano: Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, Volume 7, Issue 2, 2017. p. 81.
https://www.researchgate.net/publication/312597416_Why_a_Right_to_Explanation_of_Automated_Decision-Making_Does_Not_Exist_in_the_General_Data_Protection_Regulation; letöltés: 2022.06.11.

Az érintett hozzáférési jogából vezethető le a 15. cikk (1) rendelkezése, amely szerint „*az érintett jogosult arra, hogy az adatkezelőtől ... a következő információkhoz hozzáférést kapjon:*

...

h) ... automatizált döntéshozatal ténye, ideértve a profilalkotást is, valamint legalább ezekben az esetekben az alkalmazott logikára és arra vonatkozó érthető információk, hogy az ilyen adatkezelés milyen jelentőséggel bír, és az érintettre nézve milyen várható következményekkel jár.”

A 22. cikk kifejezetten az automatizált döntéshozatalról és a profilalkotásról szól, s ennyiben kifejezetten a mesterséges intelligencia alkalmazására, ennek törvényességi garanciáinak megalkotására irányul.

(3) „... az adatkezelő köteles megfelelő intézkedéseket tenni az érintett jogainak, szabadságainak és jogos érdekeinek védelme érdekében, ideértve az érintettnek legalább azt a jogát, hogy az adatkezelő részéről emberi beavatkozást kérjen, álláspontját kifejezze, és a döntéssel szemben kifogást nyújtson be.”

Láthatjuk, hogy az érintettet érdekeit szolgáló garanciák között a magyarázathoz való jog nem szerepel. Bizonyos védelmet ezek jelentenek, de nem adnak biztosítékot a rendszerek transzparenciájának növelésére.

Dologi jellegű kockázatok

Az intelligens rendszereket az üzleti élet számos más területén is használják – az adatkereskedelmen és a közösségi médián túl. A modern pénzügyi rendszer is digitális szolgáltatásokra épül. A bankok és a tőzsdék intelligens ügynököket használnak a pénzügyi műveletek végrehajtásához. A banki tevékenységet finom, kidolgozott jogi rendelkezések szabályozzák és védik. Az algoritmikuskereskedés – más néven a nagyfrekvenciás kereskedés – azonban viszonylag új jelenség a részvénypiacokon, és az MI-alapú algoritmusok új kihívásokat támasztanak. Az ebben az üzletágban bekövetkező pénzügyi veszteségek családok megélhetését ronghatják, és alááshatják a vállalatok prosperitását, hatalmas károkat okozva.

Az egyik leghíresebb, valószínűleg MI-algoritmusok által okozott tőzsdei incidens a 2010-es Flash Crash²⁶ a Chicagói Árutőzsdén. Az okok kivizsgálása és értelmezése még mindig folyamatban van.

A mesterséges intelligenciával támogatott pénzügyi rendszereken kívül az utakon is találkozhatunk mesterséges intelligenciával. Az autonóm járművek kockázata klasszikus etikai és jogi vitatémává nőtte ki magát. Egy hibás, tájékozatlan – vagy etikátlan – autó által okozott kár jelentős lehet, és személyi sérülést, sérülést vagy akár halált is okozhat. Ezek egyben kihívást jelentő jogi problémák is.

Ki lesz felelős egy ilyen esetben? Ez egy új terület, ahol az etikai megfontolásokat figyelembe kell venni, mielőtt jogi kereteket ki lehetne alakítani.

²⁶ BRUSH, Silla – SCHOENBERG, Tom – RING, Suzi: How a Mystery Trader with an Algorithm may have Caused the Flash Crash. Bloomberg, 2015.04.22.
<https://www.bloomberg.com/news/articles/2015-04-22/mystery-trader-armed-with-algorithms-rewrites-flash-crash-story#xj4y7vzkg>; letöltés: 2022.08.02.

A MESTERSÉGES INTELLIGENCIA ÁLTAL OKOZOTT KÁROKÉRT VALÓ FELELŐSSÉG ETIKAI ÉS JOGI ALAPJAI

„*A robot nem személy,
és még sokáig nem is lesz az.*”

Amikor eljön az öntudatos mesterséges intelligencia, vagy az erős mesterséges intelligencia, az emberfeletti mesterséges intelligencia kora – ha egyáltalán eljön –, akkor újra kell gondolnunk ezt a kijelentést, de még nem. Ezért az a fajta gyenge MI, amivel rendelkezünk, nyilvánvalóan nem lehet semmilyen jogviszony alanya, hiszen nem rendelkezik jogképességgel. A mesterséges intelligenciával szemben támasztott etikai és jogi követelmények egyaránt vonatkoznak tehát az intelligens ágenssel kapcsolatban álló jogképes személyekre, nevezetesen:

- a fejlesztőre;
- a szolgáltatóra;
- a felhasználóra.

A mesterséges intelligenciával kapcsolatos felelősség etikai alapjai

A mesterséges intelligencia fejlesztésével és működtetésével kapcsolatos etikai kritériumok közül az *átláthatóság* az első és a legfontosabb. Ez azt jelenti, hogy elkerülhetetlen a szoftverfejlesztők által használt algoritmusok harmonizálása. Bár az algoritmusok sok vállalkozás legjobban őrzött és legbizalmasabb titkai, fel kell hagyniuk az ellenőrizetlen mesterséges intelligencia használatával.

A mesterséges intelligenciára vonatkozó megkövetelt keretszabályozásról szóló tudományos, filozófiai és jogi vitákban sok további elvárás van előtérben. Ezek többsége olyan elvont elv, hogy a pontos jelentés meghatározásához kiterjedt érvelésre és értelmezésre lesz szükség.

A brit parlament felsőháza – a Lordok Háza – 2018-ban etikai kódexet²⁷ dolgozott ki az Egyesült Királyság MI-törvényéhez. Öt irányadó alapelv képezi a készülő szabályozás fundamentumát a következők szerint:

1. a mesterséges intelligenciát a *közjó* és az emberiség javára kell fejleszteni;
2. a mesterséges intelligenciának az *érthetőség* és a *méltányosság* elvei alapján kell működnie;
3. a mesterséges intelligencia nem használható az egyének, családok vagy közösségek *adattvédelmi* jogainak vagy *magánszférájának* csorbítására;

²⁷ UK can lead the way on ethical AI, says Lords Committee. UK Parliament, 2018.04.16.
<https://www.parliament.uk/external/committees/lords-select/ai-committee/news/2018/ai-report-published/>;
letöltés: 2022.05.31.

4. minden állampolgárnak joga kell legyen ahhoz, hogy a mesterséges intelligencia mellett mentálisan, érzelmileg és gazdaságilag is *boldogulni* tudjon;
5. soha nem szabad a mesterséges intelligenciát felruházni azzal az autonóm hatalommal, hogy embereknek ártson, embereket elpusztítson vagy megtévesszen.

Az erkölcsi normák jogi intézményekbe való átültetése nem triviális folyamat, s így nincs egyértelmű eredménye. Az uniós jog a morális mesterséges intelligencia fogalmát a megbízható mesterséges intelligencia fogalmával helyettesítette, és olyan kritériumokat rendelt hozzá, amelyek immár jogilag is értelmezhetőek.

Megbízható mesterséges intelligencia

A megbízható mesterséges intelligenciának²⁸ három alapelve van, amelyeknek a rendszer teljes életciklusa során meg kell felelnie:

- törvényesség;
- etikus viselkedés;
- robusztusság.

Vagyis a megbízható mesterséges intelligenciának jogszerűnek kell lennie, betartva az összes alkalmazandó törvényt és rendeletet, etikusnak kell lennie, biztosítva az etikai elvek és értékek betartását, és robusztusnak kell lennie mind technikai, mind társadalmi szempontból, mivel az MI-rendszerek még jó szándékkal is okozhatnak nem szándékos károkat.

Az alapelveket hét további konkrét követelménnyé alakítjuk át a megbízható mesterséges intelligencia megvalósítása érdekében:

1. emberi közreműködés és felügyelet;
2. műszaki megbízhatóság és biztonság;
3. adatvédelem és adatkezelés;
4. átláthatóság;
5. sokszínűség, megkülönböztetésmentesség és méltányosság;
6. társadalmi és környezeti jólét;
7. elszámoltathatóság.

A megbízható mesterséges intelligencia megfogalmazott koncepciója az alapvető jogok európai doktrínáján és az ennek megfelelő etikai elveken alapul, amelyek a mesterséges intelligencia kontextusában alapvető fontosságúak.

²⁸ Ethics Guidelines for Trustworthy AI (EGTAI). Az Európai Bizottság által 2018 júniusában létrehozott, mesterséges intelligenciával foglalkozó magas szintű független szakértői csoport megbízható mesterséges intelligenciára vonatkozó etikai iránymutatása. Európai Bizottság, Brüsszel, 2019.04.08. p. 5. <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-hu/format-PDF>; letöltés: 2022.09.04.

Az etikai kódex²⁹ kiemelte az alábbi elveket:

- emberi autonómia tiszteletben tartása;
- károkozás megelőzése;
- méltányosság;
- megmagyarázhatóság.

A megmagyarázhatóság a legtöbb problémát és fejtörést okozó elvárás az MI-fejlesztők számára, bár ez kulcsfontosságú tényező a felhasználók MI-rendszerekbe vetett bizalmának kialakításához és fenntartásához. Az MI fejlesztési és képzési folyamatainak átláthatónak kell lenniük, az MI-rendszerek képességeit és célját nyíltan kell kommunikálni, és a döntéseket a lehető legnagyobb mértékben meg kell magyarázni a közvetlenül és a közvetve érintetteknek.

A vonatkozó alapjogoknak az alapvető etikai elvekkel összefüggésben garantálniuk kell:

- az emberi méltóság tiszteletben tartását;
- az egyéni szabadságot;
- a jogállamiságot;
- a demokráciát;
- az egyenlőséget, a szolidaritást és a megkülönböztetésmentességet;
- az állampolgári jogok teljes körű érvényesülését.

Az MI-képességekkel ellátott termékeket és szolgáltatásokat nyújtó fejlesztőknek és szolgáltatóknak tartózkodniuk kell minden olyan gyakorlattól és technológiai intézkedéstől, amelyek sértik ezeket az alapvető jogi és etikai értékeket.

Ki a felelős? Ki a jogosult? Az MI jogalanyisága

A személy mint minden jog végső alanya

Valamennyi jogi attribútumnak (felelősség, tulajdonszerzési képesség, személyiségi jogok és ezek védelme stb.) szükséges előfeltétele a jogalanyiság. A jogalanyiság azt jelenti, hogy annak hordozója *SZEMÉLY*. A jog radarerőjén akkor válik valaki láthatóvá, tanulmányozhatóvá, ha a jogrendszer elismeri személynek, jogalanynak.

Mit jelent ez a mesterséges intelligencia felelősségére, személyiségi jogaira és egyéb jogi tulajdonságaira nézve?

Azt, hogy mindaddig, amíg a jogrendszer ki nem nyilvánítja, hogy az MI SZEMÉLY-nek, jogalanynak tekintendő, addig az összes további kérdés értelmetlen. Addig ugyanis a legkifinomultabb MI-rendszer is „csak egy gép”, amelynek helyes vagy hibás működéséért a fejlesztője és/vagy a használója – aki a dolgok jelen állása szerint természetes vagy jogi személy lehet – viseli a felelősséget.

²⁹ Ethics Guidelines for Trustworthy AI (EGTAI), p. 12.

Persze e körben könnyen belebotolhatunk a vétkességi alakzaton túlmenő objektív (lásd még: veszélyes üzemi) felelősség problémájába. Egy MI-alapú rendszer fejlesztése, üzemeltetése minden további nélkül tekinthető olyan veszélyes tevékenységnek, ami az objektív felelősség alkalmazását megalapozza.

Egy technológiai-filozófiai közbevetés

Igen messze vagyunk az olyan MI-rendszerek megalkotásától, amelyek valóban személyként, személyiségként viselkedhetnének, valamint a természetes személyek módjára viselnének felelősséget a „magatartásukért”.

Az MI nem is „magatartást tanúsít”, hanem funkcionál, *működik*. Működik – jól vagy rosszul. Utóbbi esetben akár még kárt is okozhat. Ezért a kárért azonban *Robot* nem lehet felelősségre vonni. Egyszerűen nem alkalmas alanya a felelősség alapját jelentő vétkesség kifejezésének, és így a jogkövetkezmények érvényesítésének.

Távol vagyunk még attól, hogy az MI tudatában legyen annak, hogy „*tőle*” milyen szabálykövető magatartást vár el a társadalom, és ezt a jogi normák rendszere hogyan érvényesíti, szankcionálja. Mégpedig azért, mert valójában az önmagára ismerő, a kontextusokat felismerő MI létrehozásától vagyunk távol.

Egy példa: tekintsük tudottnak és bizonyítottnak, hogy a DeepBlue rendszer olyan kiválóan sakkozik, hogy a játékereje összemérhető (volt 1996-ban) Garry Kaszparov sakkvilágbajnok tudásával. Ez azonban nem jelenti azt, hogy a sakkprogramot futtató IBM-számítógép mint hardver, valamint a rajta futó operációs rendszer és alkalmazások „tudták volna”, hogy ez most egy sakkturna. Azt is kizártnak tekinthetjük, hogy valamiféle tudatos – netán *érzelmi* – viszonyulása lett volna a helyzethez, amelyben működött. Biztosra vehetjük, hogy a rendszer nem „*izgult*” játék közben. *Feszültséget* is legfeljebb elektromosságtani értelemben „*érezelt*”.

Végső soron a saját emberi mentális képességeinkről is alig tudunk valamit. Minden kutatás és tiszteletre méltó tudományos eredmény ellenére ezen a területen, a kognitív tudományokban sok még a fehér folt. Nincs meg a végső és átfogó magyarázat arra, hogy mi tesz bennünket emberré, és mi az eredete, magyarázata és mechanizmusa a titokzatos *öntudatnak*, amely – Shakespeare-től tudjuk – „belőlünk mind gyávát csinál”. Az végképp hatalmas fiziológiai, neurológiai, biológiai és filozófiai rejtély, hogy mi tesz bennünket személyllyé, honnan származik az én- és öntudatunk.

Az MI személyiségi jogai

A kitérő után nézzük újra a jogi kérdést. Van értelme az MI személyiségi jogairól, felelősségéről beszélni? Semmiképpen, mielőtt a mesterséges intelligencia személyllyé nem válik. Hogy az mikor lesz? Nem tudható, de feltételezhetően nem a közeli jövőben.

A természetes személy érzékeli, érti, beilleszti a saját *öntudatvezérelt* világába a társadalomban kialakított normarendszert, a státuszát megadó kereteket, s ennek eredményeként a vele szemben alkalmazott jogi felelősséget is. Ez utóbbit nem ritkán hátrányként és korlátozásként éli meg.

Ez a hátrány lehet cselekvési szabadságtól és/vagy anyagi javaktól történő megfosztás, azok felhasználásában, élvezetében beállott korlátozás. Ez némi „áttétellel” bizonyos értelemben a jogi személyre átvihető, bár az „öntudat” fogalmát a jogi személytől is megtagadom. A természetes személy tud *szerven* (testileg és lelkileg), amit a jogi személyről nem mondhatunk el.

Az „öntudat” működésének termékeként, ennek következtében a természetes személynek van *lelkiismerete*, ami a joghátrány elmarasztaló voltának felismeréséhez elengedhetetlenül szükséges. Ez valójában már a jogi személynél sem áll fenn, és nem létezik az a technológia, aminek segítségével ezt a finom és bonyolult, nehezen megfogható lelki jelenséget bele lehetne oltani egy mesterségesen létrehozott, digitális áramkörökből alkotott lény „személyiségébe”.

Egyáltalán milyen szankciót lehetne egy MI-vel szemben megállapítani és alkalmazni? Ha rosszul működik, ki lehet cserélni valamely alkatrészét, lehet javítani, változtatni a programkódon, ad abszurdum ki lehet kapcsolni ideiglenesen vagy véglegesen, akár szét is lehet szerelni. Vajon el lehet-e érni valaha, hogy *Robi*, a *mesterséges intelligencia* felfogja, hogy őt vétkes magatartása miatt büntetés – igazságos vagy méltánytalan hátrány – éri?

Ki ne emlékezne a *Star Wars Saga* feltűnő külsejű, ügyefogyottan csetlő-botló, ügybuzgó, de kicsit félénk – tehát valódi emberi jellemvonásokkal rendelkező – protokoldroidjára. *C3PO*-t – aki kritikus helyzetekben attól retteg, hogy kikapcsolják, törlik a memóriáját, és ezzel elveszíti a személyiségét – még nem építették meg, és felettébb kétséges, hogy a ma élő nemzedékek életében erre sor kerülne. Nem nagyon valószínű, hogy az értelemből, érzelmekből, és ki tudja még milyen komponensekből összeálló emberi személyiséget egyhamar sikerül gépi alapon reprodukálni. Ez egyesek számára – nevezhetjük őket *gyanakvó technopesszimistáknak* – kicsit talán megnyugtatóan hangzik.

Ha jogalanyiságról, személyiségről, következőképp személyiségi jogokról, felelősségről és mondjuk a tulajdonszerzés képességéről, a jog által elismert lehetőségéről nem beszélhetünk, akkor miről van értelme gondolkodni az MI-jog világában? Nyilvánvalóan a mostani – határozottan „*gépnek (dolognak) tekintett*” – MI-rendszerek esetében a fejlesztő és az üzemeltető vonatkozásában merülhetnek fel mindazok a kérdések, amelyek jogi megfontolást érdemelnek.

Az MI működéséért a felelősséget a fejlesztő és/vagy a felhasználó viseli. Az MI működése következtében előállott értéket (tulajdont) szintén az említett ágensek szerzik meg. Kissé talán sántít a párhuzam, de megfelelő jogi analógiát keresve nehéz másként gondolni rá és másként kezelni: az MI egyelőre olyan fejlettségi szinten áll, hogy legfeljebb öntudatlan (rab)szolgának tekinthető, s amint a rabszolga az urának, a tulajdonosának szerzett, épp úgy *Robi* is a felhasználójának szerez, neki alkot. Ha *Robi* kárt okoz, akkor pedig szintén a gazdája viseli a felelősséget.

KONKLÚZIÓ

A mesterséges intelligencia fejlesztése – minden látványos eredmény ellenére – még mindig korai fázisában van, és még csak a kezdetén vagyunk egy hosszú útnak. A technológián kívül a mesterséges intelligenciával kapcsolatos jogi és etikai kérdések vizsgálata során is csak az első lépéseket tettük meg.

A megfelelő szabályozási keretek kidolgozásán igen sok múlhat. Nem szabad, hogy a jog hátráltassa a műszaki innovációt, a technológiai fejlődést, sőt kívánatos, hogy amennyire csak lehet, előmozdítsa azt. Ugyanakkor állandóan szem előtt kell tartania a lehetséges kockázatokat, és fenn kell tartania azokat az alkotmányos értékeket, amelyeken a demokratikus jogállamok jogrendje, szabadsága, stabilitása és anyagi jóléte alapul. A mesterséges intelligencia éppúgy lehet a szellemi és a fizikai szabadság eszköze, mint a legsötétebb elnyomásé. Sok tényezők múlik, hogy az inga melyik irányba lendül ki, és nyilvánvaló, hogy folyamatos odafigyeléssel és cselekvéssel lehet fenntartani a demokratikus, alkotmányos társadalmak jogbiztonságát – akár a mesterséges intelligencia által, akár ellenében.

A jogi keretek, a szabályozási elvek és intézményes kontrollok fogják meghatározni, hogy melyik forgatókönyv valósul meg. Ezért a szoftverfejlesztők, tudásmérnökök és jogászok vitája és együttműködése feltétlenül szükséges a mesterséges intelligencia fejlesztése során.

IRODALOMJEGYZÉK

Application of AI. javaTpoint, 2022.

<https://www.javatpoint.com/application-of-ai>; letöltés: 2022.08.11.

Artificial Intelligence Act (AIA).

Javaslat Az Európai Parlament és a Tanács Rendelete a mesterséges intelligenciára vonatkozó harmonizált jogszabályok (A Mesterséges Intelligenciáról szóló jogszabály) megállapításáról és egyes uniós jogalkotási aktusok módosításáról. Európai Bizottság, Brüsszel, 2021.04.21.

<https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52021PC0206&from=FR>; letöltés: 2022.09.04.

Artificial Intelligence Market by Offering (Hardware, Software, Services), Technology (Machine Learning, Natural Language Processing), Deployment Mode, Organization Size, Business Function (Law, Security), Vertical and Region – Global forecast to 2027. MarketResearch, August 2022.

<https://www.marketresearch.com/MarketsandMarkets-v3719/Artificial-Intelligence-Offering-Hardware-Software-32075223/>; letöltés: 2022.07.21.

BERNHARDT, Chris: Turing's Vision: The Birth of Computer Science.

The MIT Press, Cambridge, 2016.

BRUSH, Silla – SCHOENBERG, Tom – RING, Suzi: How a Mystery Trader with an Algorithm may have Caused the Flash Crash. Bloomberg, 2015.04.22.

<https://www.bloomberg.com/news/articles/2015-04-22/mystery-trader-armed-with-algorithms-rewrites-flash-crash-story#xj4y7vzkg>; letöltés: 2022.08.02.

- CARLSMITH, Joseph: Is power-seeking AI an existential risk? Open Philanthropy, April 2021.
<https://docs.google.com/document/d/1smaI1lagHHcrhoi6ohdq3TYIZv0eNWWZMPEy8C8byYg/edit#heading=h.pwdbumje5w8r>; letöltés: 2022.08.12.
- CHAN, Rosalie: The Cambridge Analytica whistleblower explains how the firm used Facebook data to sway elections. Insider, 2019.10.05.
<https://www.businessinsider.com/cambridge-analytica-whistleblower-christopher-wylie-facebook-data-2019-10>; letöltés: 2022.07.04.
- COLE, Samantha: People Are Using AI to Create Fake Porn of Their Friends and Classmates. VICE, 2018.01.26.
<https://www.vice.com/en/article/ev5eba/ai-fake-porn-of-friends-deepfakes>; letöltés: 2022.08.12.
- COLE, Samantha: We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now. A user-friendly application has resulted in an explosion of convincing face-swap porn. VICE, 2018.01.24.
<https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley>; letöltés: 2022.08.12.
- CUSTERS, Bart – FOSCH-VILLARONGA, Eduard: Humanizing Machines: Introduction and Overview. In: CUSTERS, Bart – FOSCH-VILLARONGA, Eduard (szerk.): Law and Artificial Intelligence. Regulating AI and Applying AI in Legal Practice. T.M.C. Asser Press, The Hague, 2022. pp. 3–28.
- DownAlbum. Chrome Stores.
<https://chromestores.com/extension/downalbum-extension/>; letöltés: 2022.08.25.
- Ethics Guidelines for Trustworthy AI (EGTAI).
Az Európai Bizottság által 2018 júniusában létrehozott, mesterséges intelligenciával foglalkozó magas szintű független szakértői csoport megbízható mesterséges intelligenciára vonatkozó etikai iránymutatása. Európai Bizottság, Brüsszel, 2019.04.08.
<https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-hu/format-PDF>; letöltés: 2022.09.04.
- Instagram Scraper. APIFY.
<https://apify.com/apify/instagram-scraper#need-something-different-but-still-customizable>; letöltés: 2022.08.23.
- KOZLOWSKA, Hanna: The Cambridge Analytica scandal affected nearly 40 million more people than we thought. Quartz, 2018.04.04.
<https://qz.com/1245049/the-cambridge-analytica-scandal-affected-87-million-people-facebook-says>; letöltés: 2022.05.30.
- MARR, Bernard: What are the Four Types of AI? Bernard Marr, 2021.
<https://bernardmarr.com/what-are-the-four-types-of-ai/>; letöltés: 2022.09.01.
- MASNICK, Mike: No, A ‘Supercomputer’ did NOT Pass the Turing Test for the First Time and Everyone Should Know Better. techdirt, 2014.06.09.
<https://www.techdirt.com/2014/06/09/no-supercomputer-did-not-pass-turing-test-first-time-everyone-should-know-better/>; letöltés: 2022.05.17.
- MCCARTHY, John – MINSKY, Marvin L. – ROCHESTER, Nathaniel – SHANNON, Claude E.: A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. August 31, 1955.
<http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>; letöltés: 2022.06.10.

- RUSSELL, Stuart J. – NORVIG, Peter: *Artificial Intelligence: A modern approach*. Pearson, Hoboken, New Jersey, 2021.
- SCHWARTZ, Reva – VASSILEV, Apostol – GREENE, Kristen K. – PERINE, Lori – BURT, Andrew – HALL, Patrick: *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*. NIST Special Publication 1270. NIST, 2022.03.15.
https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=934464; letöltés: 2022.06.25.
- SUWAJANAKORN, Supasorn – SEITZ, Steven M. – KEMELMACHER-SHLIZERMAN, Ira: *Synthesizing Obama: learning lip sync from audio*. *ACM Transactions on Graphics*, Volume 36, Issue 4, 20 July 2017. pp. 1–13.
<https://dl.acm.org/doi/10.1145/3072959.3073640>; letöltés: 2022.07.24.
- TAN, Jonathan Ming En: *Non-deterministic artificial intelligence systems and the future of the law on unilateral mistakes in Singapore*. *Singapore Academy of Law Journal*, 34 SAclJ 91, 2022. pp. 91–124.
<https://journalsonline.academyPublishing.org.sg/Journals/Singapore-Academy-of-Law-Journal/Current-Issue/ct/eFirstSALPDFJournalView/mid/494/ArticleId/1732/Citation/JournalsOnlinePDF>; letöltés: 2022.07.11.
- TOTSCHNIG, Wolfhart: *Fully Autonomous AI*. *Science and Engineering Ethics*, Volume 26, 2020. pp. 2473–2485.
<https://doi.org/10.1007/s11948-020-00243-z>; letöltés: 2022.06.05.
- Types of Artificial Intelligence*. javaTpoint, 2022.
<https://www.javatpoint.com/types-of-artificial-intelligence>; letöltés: 2022.08.11.
- UK can lead the way on ethical AI, says Lords Committee*. UK Parliament, 2018.04.16.
<https://www.parliament.uk/external/committees/lords-select/ai-committee/news/2018/ai-report-published/>; letöltés: 2022.05.31.
- WACHTER, Sandra – MITTELSTADT, Brent – FLORIDI, Luciano: *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*. *International Data Privacy Law*, Volume 7, Issue 2, 2017. pp. 76–99.
https://www.researchgate.net/publication/312597416_Why_a_Right_to_Explanation_of_Automated_Decision-Making_Does_Not_Exist_in_the_General_Data_Protection_Regulation; letöltés: 2022.06.11.
- What Are The 3 Types Of AI?* Deccan Herald, 2020.06.24.
<https://www.deccanherald.com/brandspot/pr-spot/what-are-the-3-types-of-ai-853275.html>; letöltés: 2022.08.10.