



NUS Law Working Paper 2020/025

A mesterséges intelligencia és a jogi személyiség határai

Simon Chesterman

chesterman@nus.edu.sg

[2020. augusztus]

Ez a tanulmány ingyenesen letölthető a Szingapúri Nemzeti Egyetem Jogi Karának munkadokumentum-sorozatának indexe: <http://law.nus.edu.sg/wps/>.

© A szerzői jog az egyes munkadokumentumok szerzőjét vagy szerzőit illeti meg. E dokumentum egyetlen része sem tehető közzé, nyomtatható újra, vagy reprodukálható bármilyen formában a dokumentum szerzőjének vagy szerzőinek engedélye nélkül.

Megjegyzés: Az egyes tanulmányokban kifejtett nézetek a tanulmány szerzőjének vagy szerzőinek nézeteit tükrözik. Nem feltétlenül képviselik vagy tükrözik a Szingapúri Nemzeti Egyetem nézeteit.

Az elektronikus kiadványra való hivatkozásokat a következő módon kell megtenni: Szerző, "Cím", NUS Law Working Paper Series, "Paper Number", a közzététel hónapja és éve, <http://law.nus.edu.sg/wps>. Például: Chan, Bala, "A Legal History of Asia," NUS Law Working Paper 2014/001, január 2014, www.law.nus.edu.sg/wps/

MESTERSÉGES INTELLIGENCIA ÉS A JOGI SZEMÉLYISÉG HATÁRAI

SIMON CHESTERMAN*

Összefoglalva, ahogy a mesterséges intelligencia (MI) rendszerek egyre kifinomultabbá válnak és egyre nagyobb szerepet játszanak a társadalomban, egyre inkább megerősödnek azok az érvek, amelyek szerint jogi személyiséggel kellene rendelkezniük. Az érveket jellemzően instrumentális szempontból fogalmazzák meg, és olyan jogi személyekkel hasonlítják össze, mint a társaságok. Ezen érvek implicit módon, illetve az illusztrációkban és példákban kifejezetten az az elképzelés jelenik meg, hogy amint az AI-rendszerek megközelítik az embertől való megkülönböztethetlenség határát, a természetes személyekhez hasonló jogállásra kell őket feljogosítani. Ez a cikk azt állítja, hogy bár a legtöbb jogrendszer létrehozhatná a jogi személyek új kategóriáját, az ilyen érvek nem elegendőek annak bizonyítására, hogy ezt meg kellene tenniük.

Kulcsszavak: mesterséges intelligencia, jogi személyiség, természetes személy, jogi személy, társaság, társasági fátyol, szerzői jog, szabadalom.

2021-ben a Bank of England az új ötvenfontos bankjegy kibocsátásával befejezi a papírról a polimerre való átállást. A nyilvános kiválasztási folyamat során közel negyedmillióan jelölték az új bankjegy arculatát; a júliusban bejelentett végső döntés 2019, szerint Alan Turing lesz a címlapon. Turing a második világháború alatti kódfejtés hőse volt. Emellett segítette megalapozni a számítástechnika tudományágát, megteremtve annak alapjait, amit ma mesterséges intelligenciának (AI) nevezünk.¹ Talán a legismertebb

* BA/LLB (Hons) (Melbourne), DPhil (Oxford). A Szingapúri Nemzeti Egyetem jogi karának dékánja és provostori székfoglaló professzora. Köszönettel tartozom Damian Chalmersnek, Hu Yingnek, Kenneth Khoonak, Mark Limnek, Lin Lin Lin-nek, Daniel Sengnek, David Tan-nak, Tan Zhong Xingnek, Hans Tjio-nak, Umakanth Varottilnak, Wong Chee Leongnak és két névtelen bírálónak a szöveg korábbi változataihoz fűzött észrevételeikért. Felbecsülhetetlen értékű kutatási segítséget nyújtott Violet Huang, Eugene Lau, Ong Kye Jing és Yap Jia Qing. A hibák és kihagyások kizárólag a szerzőt terhelik.

¹ A mesterséges intelligencia meghatározására tett kísérletekről lásd: SJ Russell és P Norvig, *Artificial Intelligence: A Modern Approach* (3rd edn, Prentice Hall 2010) 1-5. Négy átfogó megközelítést

különböztethetünk meg: emberi módon cselekedni (Turing-teszt), emberi módon gondolkodni (kognitív viselkedés modellezése), racionálisan gondolkodni (a logikai hagyományra építve), és

a hozzájárulás azonban a névadó teszt arra, hogy mikor valósult meg az igazi "intelligencia".

Turing egy 1950-ben népszerű társasjátékról mintázta. Egy férfi és egy nő külön szobában ül, és írásban válaszol kérdésekre; a többi résztvevőnek ki kell találnia, hogy ki melyik választ adta. Turing felvetette, hogy hasonló "utánzójátékot" lehetne játszani számítógéppel is. Ha egy gép képes elhíttetni az emberekkel, hogy ember, akkor joggal mondhatjuk, hogy intelligens.²

A korai sikerek ezen a vonalon az 1960-as években születtek az olyan programokkal, mint az Eliza. A felhasználóknak azt mondták, hogy Eliza egy pszichoterapeuta, aki a számítógépbe beírt szavakkal kommunikál. Valójában "ő" egy algoritmus volt, amely egy egyszerű listafeldolgozó nyelvet használt. Ha a felhasználó beírt egy felismert kifejezést, az kérdésként lett átfórmálva. Tehát miután beírta, hogy "depressziós vagyok", Eliza azt válaszolhatta, hogy "Miért mondd, hogy depressziós vagy?". Ha nem ismerte fel a kifejezést, akkor a program valami általánosabbat ajánlott, például: "Elmondanád ezt bővebben?". Néhány felhasználó még akkor is ragaszkodott ahhoz, hogy Eliza "megértette" őket, amikor elmondták nekik, hogyan működik a program.³

A társasjátékoktól eltekintve, miért számítana, hogy egy számítógép "intelligens"? A Turing-tesztet évtizedekig inkább azzal a kérdéssel hozták összefüggésbe, hogy lehetséges-e maga a mesterséges intelligencia, mintsem egy ilyen entitás jogi státuszával. Mégis gyakran hivatkoznak rá a mesterséges intelligencia jogi személyiségéről szóló vitákban, Lawrence Solum 1992-es, korszakalkotó cikkétől kezdve.⁴ Bár technikai értelemben már nem tekintik a modern mesterséges intelligencia komoly mércéjének, a Turing-teszt mint trópus hosszú életűsége rámutat a személyiséggel kapcsolatos viták olyan feszültségére, amelyet gyakran figyelmen kívül hagynak.

Ahogy a mesterséges intelligencia rendszerek egyre kifinomultabbá válnak és egyre nagyobb szerepet játszanak a társadalomban, legalább két különálló ok van arra, hogy miért lehetne őket a törvény előtt személyként elismerni. Az első az, hogy legyen kit hibáztatni, ha a dolgok rosszul mennek. Ez a válasz a gyorsaságuk, autonómiájuk és átláthatatlanságuk által okozott lehetséges elszámoltathatósági hiányosságokra.⁵ A személyiség elismerésének második oka azonban az, hogy biztosítsák, hogy legyen kit jutalmazni.

raciónális cselekszik (a Russell és Norvig által kedvelt racionális-agens megközelítés, mivel nem függ az emberi megismerés sajátos értelmezésétől vagy a racionális gondolkodás kimerítő modelljétől).

² AM Turing, "Computing Machinery and Intelligence" (Számítógépek és intelligencia) (1950) Mind 59433.

³ RS Wallace, "The Anatomy of ALICE" in R Epstein, G Roberts, and G Beber (szerk.), *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer* (Springer 2009) 184-85.

⁴ LB Solum, "Legal Personhood for Artificial Intelligences" (1992) North Carolina L Rev 1235-37. Solum maga Christopher Stone-nak tulajdonítja, hogy két évtizeddel korábban egy lábjegyzetben először vetette fel ezt a

lehetőséget: CD Stone, "Should Trees Have Standing? Towards Legal Rights for Natural Objects" (1972) Southern California L Rev n26.456

⁵ Lásd S Chesterman, "Artificial Intelligence and the Problem of Autonomy" (2020) 1 Notre Dame J Emerging Technologies 210; S Chesterman, "Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity" (2021) AJCL megjelenés alatt.

amikor a dolgok jól mennek. Egyre több szakirodalom foglalkozik például az AI-rendszerek által létrehozott szellemi tulajdonjogokkal.

E vitákban az a feszültség, hogy a személyiséget instrumentális vagy inherens okokból adják-e meg. Az érveket jellemzően instrumentális szempontok alapján fogalmazzák meg, a leggyakoribb mesterséges jogi személyhez, a vállalathoz hasonlítva. Ugyanakkor számos ilyen érven implicit módon, vagy az illusztrációkban és példákban kifejezetten az az elképzelés jelenik meg, hogy amint a mesterséges intelligenciával működő rendszerek megközelítik az embertől való megkülönböztethetlenség határát - vagyis amikor átmennek a Turing-teszten -, a természetes személyekhez hasonló státuszra kell őket feljogosítani.

A közelmúltig az ilyen érvek mind spekulatívak voltak. Aztán 2017-ben Szaúd-Arábia "állampolgárságot" adott a humanoid robotnak, Sophiának⁶, és egy hétéves kisfiú személyiségével rendelkező online rendszer "tartózkodási engedélyt" kapott Tokióban.⁷ Ezek trükkök voltak - Sophia például lényegében egy chatrobot arccal.⁸ Ugyanebben az évben azonban az Európai Parlament állásfoglalást fogadott el, amelyben felszólította a Bizottságot, hogy fontolja meg "a robotok sajátos jogi státuszának létrehozását hosszú távon, hogy legalább a legfejlettebb autonóm robotok elektronikus személy státuszát lehessen megállapítani, akik felelősek az általuk okozott károk megtérítéséért, és esetleg elektronikus személyiséget lehessen alkalmazni azokra az esetekre, amikor a robotok autonóm döntéseket hoznak vagy más módon önállóan lépnek kapcsolatba harmadik felekkel."⁹

Ez a cikk a legközvetlenebb kihívással kezdi, vagyis azzal, hogy a jogi személyiség valamilyen formája betöltene-e egy felelősségi űrt, vagy más módon előnyös lenne-e a jogrendszer számára. A társaságok és más mesterséges jogi személyek története alapján nem tűnik kétségesnek, hogy a legtöbb jogrendszer *képes lenne a* mesterséges intelligenciarendszereknek egyfajta személyiséget biztosítani; az érdekesebb kérdés az, hogy kell-e, és milyen tartalmú lehet ez a személyiség.

A második szakasz ezután a természetes személyekkel való analógia felé fordul. Magától értetődőnek tűnhet, hogy egy gép soha nem lehet természetes személy. Évszázadokon keresztül azonban a rabszolgákat és a nőket sem ismerték el teljes jogú személyként. Ha a Turing-tesztet a logikus, *Blade Runner*-féle következtetésig visszük, lehetséges, hogy az emberektől valóban megkülönböztethetetlen mesterséges intelligencia-rendszerek egy napon ugyanezt a státuszt követelhetik. Bár a "robotok jogairól" szóló viták jelenleg a következő témákra korlátozódnak

⁶ O Cuthbert, "Saudi Arabia Becomes First Country to Grant Citizenship to a Robot", *Arab News* (2017. október 26.).

⁷ A Cuthbertson, "Artificial Intelligence "Boy" Shibuya Mirai Becomes World's First AI Bot to Be Granted Residency", *Newsweek*, 2017. november 6.

⁸ D Gershgorn, "Inside the Mechanical Brain of the World's First Robot Citizen", *Quartz* (2017. november 12.).

⁹ Az Európai Parlament állásfoglalása a Bizottságnak szóló ajánlásokkal a robotikára vonatkozó polgári jogi szabályokról (2015/2103(INL)) (Európai Parlament, 162017. február), 59. bekezdés f) pont.

a diskurzus perifériájára szorult, ez a lehetőség implicit módon jelen van számos érvben, amelyek mellett szólnak, hogy a mesterséges intelligencia rendszerek az általuk létrehozott szellemi tulajdonnal rendelkezzenek.

Ha komolyan vesszük, az az elképzelés, hogy a mesterséges intelligencia rendszerek az emberekkel egyenrangúak lehetnek, egy harmadik okot is felvet a személyiségről való gondolkodásra. Ha ugyanis az egyenlőség egyszer megvalósul, nincs okunk feltételezni, hogy az AI fejlődése megállna. Bár az általános mesterséges intelligencia egyelőre science fiction marad, megfontolásra késztet, hogy a jogi státusz alakíthatja vagy korlátozhatja-e a viselkedést, ha vagy amikor az emberiséget felülmúlják. Ha erre valaha is sor kerülne, a kérdés persze nem az lenne, hogy elismerjük-e az általános mesterséges intelligencia jogait, hanem az, hogy elismeri-e a mi jogainkat.

I. JOGI SZEMÉLYISÉG: TESTET RÚGNI?

A jogi személyiség minden jogrendszer alapvető eleme. Az a kérdés, hogy ki cselekedhet, ki lehet jogok és kötelezettségek alanya, szinte minden más kérdés előzménye. Mégis, ezen alapok alapos vizsgálata meglepő bizonytalanságot és nézeteltérést tár fel. Ennek ellenére, ahogy John Dewey 1926-ban megjegyezte, "a bíróságok és a jogalkotók ilyen egyetértés nélkül, olykor mindenféle koncepció vagy elmélet nélkül végzik munkájukat" a személyiség természetével kapcsolatban. Sőt, folytatta, az elmélethez való folyamodás "nem egyszer inkább akadályozta, mint megkönnyítette a jog vagy kötelezettség egy speciális kérdésének eldöntését."¹⁰

A gyakorlatban a jogrendszerek túlnyomó többsége a jogi személy két formáját ismeri: a természetes és a jogi személyt. A természetes személyeket azon egyszerű tény miatt ismerik el, hogy emberek.¹¹ A jogi személyek ezzel szemben olyan nem emberi entitások, amelyeket a törvény bizonyos jogokkal és kötelezettségekkel ruház fel. A társaságok és az üzleti társulások más formái a leggyakoribb példák, de számos más forma is lehetséges. Nemzeti és nemzetközi szinten vallási, kormányzati és kormányközi szervezetek is felléphetnek jogi személyként.

Beszédes, hogy ezek mind emberi szereplők összességei, bár van példa arra, hogy valóban nem emberi entitásoknak is megadják a személyiséget. Az itt említett példákon kívül

¹⁰ J Dewey, "The Historic Background of Corporate Legal Personality" (1926) Yale35 LJ 660.

¹¹ Ez természetesen feltételezi, hogy egyetértés van az "ember" fogalmának és az olyan fogalmaknak a jelentésében, mint a születés és a halál. Lásd N Naffine, "Kik a jog személyei? From Cheshire Cats to Responsible Subjects" (2003) MLR66 (2003). 346.

bevezetés, ezek közé tartoznak az indiai templomok,¹² egy folyó Új-Zélandon,¹³ és Ecuador teljes ökoszisztémája.¹⁴ Nem tűnik kétségesnek, hogy egy állam valamiféle személyiséget tulajdoníthat az olyan új entitásoknak, mint az AI-rendszerek;¹⁵ ha ez megtörténik, akkor valószínűleg más államok is elismerik majd.¹⁶

A. *A jogi személyiség elméletei*

A tudósok és a jogreformmal foglalkozó szervek már javasolták, hogy a felelősséggel kapcsolatos kérdések megoldásának megkönnyítése érdekében a mesterséges intelligenciával rendelkező rendszereket valamilyen jogi személyiséggel ruházzák fel, például az automatizált vezetési rendszerek jogalanyaként a vezető nélküli autók esetében, amelyek viselkedése nem feltétlenül a "vezetőjük" irányítása alatt áll, vagy nem a gyártó vagy a tulajdonos által kiszámítható.¹⁷ Néhány író ennél is tovább ment, és azt állította, hogy eljárásokat kell bevezetni a robotbűnözők bíróság elé állítására, és rendelkezni kell az átprogramozással vagy - szélsőséges esetben - megsemmisítéssel történő "büntetésről".¹⁸

Ezek az érvek a személyiség instrumentális megközelítését sugallják, de a jogi személy legerjedtebb formájának - a társaságnak - a tudományos magyarázatai eltérő indoklásokat mutatnak az önálló jogi személy státuszára, ami segíthet megválaszolni azt a kérdést, hogy ezt a státuszt ki kell-e terjeszteni a mesterséges intelligencia rendszerekre is.

Az aggregált elmélet, amelyet néha kontraktuális vagy szimbolista elméletnek is neveznek, azt állítja, hogy a társaság a jog által létrehozott eszköz, amely lehetővé teszi a magukat csoportként szervező természetes személyek számára, hogy ezt a szervezetet tükrözzék a más felekkel való jogi kapcsolataikban. A csoport tagjai létrehozhatnának egyéni szerződéses kapcsolatokat ezekkel a többi féllel, amelyek korlátozzák a felelősséget stb.

¹² Lásd pl. *Shiromani Gurdwara Prabandhak Committee, Amritsar kontra Shri Somnath Dass* AIR 2000 SC 1421 (India Legfelsőbb Bírósága).

¹³ Te Awa Tupua (Whanganui River Claims Settlement) Act 2017 (Új-Zéland), 14. szakasz (1) bekezdés. Ez azt követően történt, hogy a Te Urewera Nemzeti Parkot "jogi személlyé nyilvánították, [a jogi személyek minden jogával, hatáskörével, kötelességével és felelősségével]". Te Urewera Act 2014 (Új-Zéland), 11. szakasz (1) bekezdés.

¹⁴ Az Ecuadori Köztársaság alkotmánya (2008Ecuador), cikkelye 10.

¹⁵ A jogi személyiség határainak megvitatásához lásd VAI Kurki, *A Theory of Legal Personhood* (OUP 2019) 127-52.

¹⁶ Lásd pl. *Bumper Development Corp. kontra Commissioner of Police for the Metropolis* [1991] WLR1 (egy indiai templom jogi személyiségének 1362elismerése az angol jog szerint).

¹⁷ Lásd pl.: *Changing Driving Laws to Support Automated Vehicles (Policy Paper)* (Nemzeti Közlekedési Bizottság, 2018. május), 1.5. bekezdés; *Automated Vehicles: A Joint Preliminary Consultation Paper* (Law Commission, Consultation Paper No. 240; Scottish Law Commission, Discussion Paper No 2018166.), (Vö.4.107. Chesterman, "Autonomy" (n 5) 225.

¹⁸ Lásd pl. G Halleve, *Liability for Crimes Involving Artificial Intelligence Systems* (Springer 2015); Y Hu, "Robot Criminals" (2019) 52 U Michigan JL Reform. 487.

alacsonyabb költségek.¹⁹ Az elméletet kritika érte, és mindenestre a legkevésbé alkalmazható a mesterséges intelligencia rendszerekre.²⁰

A társasági személyiség fikciós és koncessziós elmélete különálló eredetű, de²¹ ugyanarra a dologra vezethető vissza: a társaságok azért rendelkeznek személyiséggel, mert a jogrendszer úgy dönt, hogy azt megadja nekik. Amint azt az Egyesült Államok Legfelsőbb Bírósága megállapította, a 1819. társaság "mesterséges lény, láthatatlan, megfoghatatlan, és csak a jog szemléletében létezik".²² A személyiséget politikai célok elérése érdekében adják meg, például a vállalkozói szellem ösztönzése érdekében, vagy azért, hogy hozzájáruljanak a jogrendszer koherenciájához és stabilitásához, például bizonyos szervezetek örökkévalósága révén. A céltudatos aspektus régebben sokkal nyilvánvalóbb volt, amikor a személyiséget kifejezetten alapító okirat vagy jogszabály útján adták meg; a huszadik század folyamán ez pusztán formalitássá vált.²³ Ezek a pozitivisták beszámolóik a személyiség elismerése terén a legjobban illeszkednek a jogalkotási és bírói gyakorlathoz, és magukban foglalhatják annak kiterjesztését a mesterséges intelligencia rendszerekre.

A realista elmélet ezzel szemben azt állítja, hogy a vállalatok nem fikciók vagy pusztán szimbólumok, hanem objektíven valós entitások, amelyeknek a jogrendszer által történő személyisége már nem létezik. Bár lehetnek tagjaik, önállóan cselekszenek, és cselekedeteik nem tulajdoníthatók e tagoknak. A legszélsőséges esetben azt állítják, hogy a társaságok nemcsak jogi, hanem erkölcsi személyek is.²⁴ Ezt az elméletet inkább a teoretikusok és a szociológusok, mint a jogalkotók és a bírák kedvelik, de a bevezetőben kiemelt feszültséget tükrözi: a jogi személyiséget nem pusztán adományozzák, hanem ki is érdemlik. A gyakorlatban azonban a személyiség tényleges elismerése a jog előtt az állam ajándéka marad.²⁵

A végeredmény talán az, hogy Dewey-nak száz évvel ezelőtt igaza volt: ""személy" azt jelenti, amit a jog jelent".²⁶ Bár a személyiség kérdése bináris, azonban - elismert vagy nem elismert - e státusz tartalma egy spektrum. Egy pillanatra félretéve azt az elképzelést, hogy egy

¹⁹ R Coase, "The Nature of the Firm" (1937) 4 *Economica* 386. Vö. V Morawetz, *A Treatise on the Law of Private Corporations* (Little, Brown 1886) (2"az a tény, hogy a vállalat valójában nem egy személy vagy egy dolog, amely különbözik az alkotó részeitől. A társaság szó nem más, mint a társasági tagok gyűjtőneve").

²⁰ N Banteka, "Artificially Intelligent Persons" (2020) *Houston* 58 L Rev forthcoming.

²¹ Dewey (10. o.) 665-69.

²² *Trustees of Dartmouth Coll. kontra Woodward*, USA (17518,1819636).

²³ Lásd pl. CE Amsler, RL Bartlett és CJ Bolton, "Thoughts of Some British Economists on Early Limited Liability and Corporate Legislation" (1981) *History* 13 of Political Economy 774; G Dari-Mattiacci *et al*, "The Emergence of the Corporate Form" (2017) 33 *J of L, Economics and Organization*. 193.

²⁴ P French, "The Corporation as a Moral Person" (1979) 16(3) American Philosophical Q 207.

²⁵ K Iwai, "Személyek, dolgok és vállalatok: The Corporate Personality Controversy and Comparative Corporate Governance" (1999) AJCL47 583; SM Watson, "The Corporate Legal Person" (2019) J19 Corporate L Studies. 137.

²⁶ Dewey (n 10) 655.

A mesterséges intelligencia rendszer személyként való elismerést érdemelhet, az államnak az elismerésről szóló döntését a szintén elismert jogoknak és köteleességeknek kell vezérelniük.

B. A jogi személyiség tartalma

A jogi személyiség jogokkal és kötelezettségekkel jár, de ezeknek nem kell azonosnak lenniük egy jogrendszerben minden személy számára. Még a természetes személyek körében is tükrözi ezt az igazságot a nők, az etnikai vagy vallási kisebbségek és más hátrányos helyzetű csoportok egyenlő jogaiért folytatott küzdelem.²⁷

Lehetőség van például arra, hogy csak jogokat adjunk kötelezettségek nélkül. Ez a megközelítés volt a jellemző a természet személyiségének megadásakor - mind elméletben, amikor először szorgalmazták, mind 1972,^{28a} gyakorlatban, például Ecuador alkotmányában.²⁹ Lehetne azzal érvelni, hogy az ilyen "személyiség" csupán egy csel, hogy elkerüljék a jogállási problémákat: lehetővé teszi az emberi személyek számára, hogy egy nem emberi jogtulajdonos nevében cselekedjenek, ahelyett, hogy saját minőségükben kellene megalapozniuk a jogállást.³⁰ Mindenesetre ez nem tűnik megfelelőnek a mesterséges intelligenciával rendelkező rendszerek személyiségének mérlegelését indokoló okok szempontjából.

Másrészt a mesterséges intelligencia jogi személyiség csak kötelezettségekkel járhat. Ez felszínesen vonzónak tűnhet, de amennyiben e kötelezettségek célja az elszámoltathatósági hiányosságok kezelése, nyilvánvaló problémák merülnének fel. A polgári jogi felelősség például jellemzően kártérítés megítéléséhez vezet, amelyet csak akkor lehet megfizetni, ha a vétkes képes a tulajdonjog birtoklására.³¹ Elképzelhetőek olyan forgatókönyvek, amelyekben ezeket a kifizetéseket egy központi alapból teljesítik, bár ez inkább a felelősség kérdésének alternatív megoldásaként javasolt kötelező biztosítási rendszerekhez hasonlítana.³² A "személyiség" csupán formalitás lenne.

²⁷ JJ Bryson, ME Diamantis és TD Grant, "A népből, a népért és a nép által: The Legal Lacuna of Synthetic Persons" (2017) 25 Artificial Intelligence and Law (Mesterséges intelligencia és jog). 280.

²⁸ Stone (n 4).

²⁹ Az Ecuadori Köztársaság alkotmánya, 71-74. cikk.

³⁰ C Rodgers, "Az ökoszisztémák védelmének új megközelítése" (2017) 19 Environmental L Rev 266. Új-Zélandon ezzel szemben a személyiséggel rendelkező környezeti jellemzők nevében történő cselekvésre létrehozott vagyonkezelőket hoztak létre.

³¹ Vö. a mesterséges intelligenciával és más újonnan megjelenő digitális technológiákkal kapcsolatos felelősség Személyiség v43a (28-Aug-20)

(A felelősséggel és az új technológiákkal foglalkozó uniós szakértői csoport, 2019) 38.

³² Lásd pl. DA Crane, KD Logue és BC Pilz, "A Survey of Legal Issues Arising from the Deployment of Autonomous and Connected Vehicles" (2017) *Michigan Telecommunications and Technology L Rev* 256-59. Vö. KS Abraham és RL Rabin, "Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era" (2019) *105 Virginia L Rev.* 127.

A társaságok esetében a személyiség jellemzően a perlési és perelhető, szerződéskötési, adóssághelyettesítési, tulajdonszerzési és büntetethetőségi képességet jelenti. Ami a jogokat illeti, folyamatos vita tárgyát képezi, hogy a természetes személyekhez hasonló alkotmányos védelmet élveznek-e a társaságok. Bár az Egyesült Államok vitathatatlanul a legtöbb védelmet biztosította a vállalati jogalanyoknak, még ott is meghúztak egy határt az olyan garanciáknál, mint az önvádolás elleni jog.³³ A jogi személyek általában kevesebb joggal rendelkeznek, mint a természetes személyek. (Hasonló a helyzet a nemzetközi jogban is, ahol az államok teljes jogú személyiséggel rendelkeznek, a nemzetközi szervezetek pedig különböző mértékben.³⁴)

1. Magánjog

Az Európai Parlament elismerte, hogy a személyiség egyik fő vonzereje a mesterséges intelligencia rendszerek számára a perelhetőség.³⁵ Ez persze feltételezi, hogy vannak értelmes elszámoltathatósági hiányosságok, amelyeket be lehet és be is kell tölteni. Ezeket a hiányosságokat gyakran túlértékelik.³⁶ Az ilyen jogorvoslattal kapcsolatos óvatosságra az ad okot, hogy még ha a személyiség megadása a mesterséges intelligenciával működő rendszereknek hiánypótló funkciót is látna el, a jelenlegi jogszabályok szerinti felelősséget a meglévő jogi személyekről hárítaná el. Sőt, ez arra ösztönözné, hogy a kockázatot az ilyen elektronikus személyekre hárítsák át, hogy a természetes és hagyományos jogi személyeket megóvják a kitétségtől.³⁷ Ez a probléma a társaságok esetében is fennáll, amelyeket arra lehet használni, hogy megvédjék a befektetőket a befektetésük fix összegén túli felelősségtől - sőt, gyakran éppen ez a célja a társasági forma használatának. A kockázat átcsoportosítását azzal indokolják, hogy az ösztönzi a beruházásokat és a vállalkozói szellemet.³⁸ A biztosítékok közé tartozik általában az a követelmény, hogy a korlátolt felelősségű jogalany nevének tartalmaznia kell ezt a státuszt a nevében ("Ltd", "LLC" stb.), valamint a társasági fátlyol korlátozott körülmények között történő áttörésének lehetősége, hogy megakadályozzák a visszaélést a korlátolt felelősséggel való visszaélést.

³³ SA Trainor, "A Comparative Analysis of a Corporation's Right Against Self-Incrimination" (1994) 18 Fordham International LJ 2139. Lásd azonban *Citizens United kontra Federal Election Commission*, 558 US 310 (US Supreme Court, 2010).

³⁴ S Chesterman, "Létezik-e az ASEAN? The Association of Southeast Asian Nations as an International Legal Person" (2008) XII Singapore Year Book of International Law (Szingapúri Nemzetközi Jogi Évkönyv) 199.

³⁵ Lásd fent n 9.

³⁶ Lásd Chesterman, "Autonomy" (5. sz.); Chesterman, "Opacity" (5. sz.).

³⁷ Bryson, Diamantis és Grant (n 27) 287.

³⁸ FH Easterbrook és DR Fischel, "Limited Liability and the Corporation" (1985) U52 Chicago L Rev. 89.

formában.³⁹ A mesterséges intelligencia rendszerek esetében hasonló fátylat áttörő mechanizmusokat lehetne kifejleszteni - bár ha egy ember manipulálná a mesterséges intelligenciát annak érdekében, hogy megvédje magát a felelősségtől, akkor az erre való képesség azt sugallhatná, hogy a szóban forgó mesterséges intelligencia rendszer nem érdemli meg a külön személyiséget.⁴⁰

A szerződéskötést időnként felvetik, mint indokot arra, hogy a mesterséges intelligencia rendszereket személyiséggel ruházzák fel.⁴¹ Az elektronikus ügynökök kötelező erejű szerződések megkötésére való felhasználása azonban aligha újdonság. A nagyfrekvenciás kereskedés például olyan algoritmusokra támaszkodik, amelyek hagyományos személyek nevében más algoritmusokkal kötnek megállapodásokat.⁴² Bár az AI-rendszerek autonómiája megkérdőjelezheti a meglévő doktrína alkalmazását az ilyen gyakorlatokra - nevezetesen, ha valami rosszul sül el, például hiba történik -, ez még mindig megoldható új jogi személyek igénybevétele nélkül.⁴³

Az adósságvállalás és a tulajdonlás szükségessége velejárója lenne a perelhető és szerződéskötési képességnek.⁴⁴ Az a lehetőség, hogy a mesterséges intelligencia rendszerek vagyont halmozhatnak fel, felveti azt a kérdést, hogy meg lehet-e adóztatni őket, vagy ha igen, hogyan. A robotok megadóztatását javasolták az automatizálás következtében várhatóan csökkenő adóalap és a munkavállalók kiszorulásának kezelésére.⁴⁵ Többek között Bill Gates is javasolta, hogy az ilyen robotokat - vagy az őket birtokló vállalatokat - meg kellene adóztatni.⁴⁶ Az ipar képviselői azzal érveltek, hogy ez negatív hatással lenne a versenyképességre, és eddig nem fogadták el.⁴⁷ Egy másik lehetőség, hogy nem a gépeket, hanem a piaci pozícióval visszaélő vállalatok helyzetét vizsgálják, és ennek keretében a nyereség agresszívebb megadóztatását vagy a következőkre vonatkozó követelményeket határozzák meg.

³⁹ D Millon, "Piercing the Corporate Veil, Financial Responsibility, and the Limits of Limited Liability" (2007) 56 Emory LJ (A vállalati fátylat áttörése, pénzügyi felelősség és a korlátozott felelősség határai) (2007). 1305.

⁴⁰ J Turner, *Robot szabályok: Regulating Artificial Intelligence* (Palgrave Macmillan 2019). 193.

⁴¹ Lásd pl. S Chopra és LF White, *A Legal Theory for Autonomous Artificial Agents* (U Michigan Press 2011). 160.

⁴² Lásd pl. T Čuk és A van Waeyenberge, "European Legal Framework for Algorithmic and High Frequency Trading (Mifid 2 and MAR) A Global Approach to Managing the Risks of the Modern Trading Paradigm" (2018) 9 European J Risk Regulation. 146.

⁴³ Lásd pl. *Quoine Pte Ltd kontra B2C2 Ltd* [2020] SGCA(I) 2 (Szingapúri Fellebbviteli Bíróság), 97-128. pont; Chesterman, "Autonomy" (n. 5) 243-44.

⁴⁴ Egyesek szerint ez a társaságok különálló jogi személyiségének legfontosabb funkciója: H Hansmann és R Kraakman, "The Essential Role of Organizational Law" (2000) 110 Yale LJ 387. Vö. H Tjio, "Lifting the Veil on Piercing the Veil" [2014] Lloyd's Maritime and Commercial Law Quarterly Elképzelhető19., hogy a tulajdonjoggal nem rendelkező mesterséges intelligenciarendszerek továbbra is a jogi eljárások bizonyos formáinak, például a végzéseknak a hatálya alá tartozhatnak, és "munkájukkal" adósságokat egyenlíthetnek ki.

⁴⁵ BA King, T Hammond és J Harrington, "Disruptive Technology: (2017) 12(2) J Strategic Innovation and Sustainability 53.

⁴⁶ KJ Delaney, 'The Robot that Takes Your Job Should Pay Taxes, says Bill Gates', *Quartz*, February 18 2017.

⁴⁷ G Prodhon, "European Parliament Calls for Robot Law, Rejects Robot Tax", *Reuters* (2017. február 16.); L Summers, "Robots Are Wealth Creators and Taxing Them Is Illogical", *Financial Times* (2017. március 6.).

elosztott részvénytulajdon.⁴⁸ Mindenesetre a mesterséges intelligencia rendszerek megadóztatása - mint például az adósságfelvétel és a tulajdonlás képessége - inkább következne belőlük, mintsem indokolná a személyiség megadásának megalapozását.⁴⁹ (A mesterséges intelligenciarendszerek alkotásaik tulajdonjogának kérdésével a következő szakaszban foglalkozunk.⁵⁰)

Az ingatlanok birtoklása mellett a mesterséges intelligencia-rendszerek feladata lehet azok kezelése is. 2014-ben például bejelentették, hogy egy hongkongi kockázati tőkebefektető cég egy Vital nevű számítógépes programot nevezett ki az igazgatótanácsába.⁵¹ Akárcsak a szaúd-arábiai kormány állampolgárságának odaítélése, ez is inkább stílus volt, mint tartalom - a hongkongi törvények értelmében a programot nem nevezték ki semmire; egy interjúban néhány évvel később az ügyvezető partner elismerte, hogy a cég Vital-t csupán megfigyelői státuszú tagként kezelte az igazgatótanácsban.⁵² Lehetséges, hogy az emberi igazgatók átruházhatnak bizonyos felelősséget egy mesterséges intelligencia rendszerre, de a legtöbb társasági jogi rendszer szerint nem mentesíthetik magukat a szervezet irányításának végső felelőssége alól.⁵³ A legtöbb jogrendszer megköveteli, hogy az igazgatók természetes személyek legyenek, bár egyes jogrendszerekben lehetséges, hogy jogi személy - jellemzően egy másik vállalat - is tagja legyen az igazgatótanácsnak.⁵⁴ Shawn Bayern ennél is tovább ment, és azzal érvelt, hogy az amerikai gazdasági társasági jog kikapuit kihasználva olyan korlátolt felelősségű társaságokat lehet létrehozni, amelyeknek egyáltalán nincsenek emberi tagjai.⁵⁵ Ez a törvény némileg kínkeserves értelmezését igényli - egy természetes személy létrehoz egy vállalatot, felvesz egy mesterséges intelligencia rendszert tagként, majd lemond⁵⁶-, de azt sugallja, hogy a jogi személyiséget hogyan lehet a jövőben átalakítani.

⁴⁸ "Why Taxing Robots Is not a Good Idea", *Economist* (2017. február 25.).

⁴⁹ Vö. L Floridi, "Robots, Jobs, Taxes, and Responsibilities" (2017) *Philosophy30 & Technology* 1.

⁵⁰ Lásd az alábbi II.B. szakaszt.

⁵¹ R Wile, "A Venture Capital Firm Just Named an Algorithm to Its Board of Directors", *Business Insider*, május 132014.

⁵² N Burridge, "AI Takes Its Place in the Boardroom", *Nikkei Asian Review*, május. 252017.

⁵³ F Möslin, "Robotok az igazgatótanácsban: (szerk.), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar 2018) 658-60.

⁵⁴ Lásd pl. 1926. évi személyek és társasági jog (PGR) (Liechtenstein), 344. cikk; 2014. évi társasági rendelet (HK), 457. szakasz. Ez 2015-ig az angol jog szerint is lehetséges volt. Lásd most Small Business, Enterprise and Employment Act 2015 (UK), section 87.

⁵⁵ S Bayern, "Of Bitcoins, Independently Wealthy Software, and the Zero-Member LLC" (2014) Northwestern 108 U L Rev 1495-500.

⁵⁶ S Bayern, "The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems" (2015) Stanford 19 Technology L Rev. 101.

2. Büntetőjog

A jogalanyiség utolsó tulajdonsága a leginkább zsigeri, és megérdemli a részletesebb kifejtést: a büntethetőség. Ha egy társasághoz hasonló jogi személyiséggel ruháznánk fel, akkor kevés okunk lenne vitatkozni azon, hogy egy mesterséges intelligencia rendszer büntetőjogi felelősségre vonható lenne-e. Amennyiben az *actus reus* és a *mens rea* megállapítható, egy⁵⁷ ilyen szervezetet pénzbírsággal lehetne sújtani, vagy el lehetne foglalni a tulajdonát; a működési engedélyt fel lehetne függeszteni vagy vissza lehetne vonni. Egyes jogrendszerekben a jogi személy ellen felszámolási végzés hozható; amennyiben ez nem lehetséges, a szervezet csődjéhez elegendő pénzbírság ugyanilyen hatású lehet. Szélsőséges esetben elképzelhető, hogy egy "robotbűnözőt" megsemmisítenek. De vajon kívánatos lenne-e ez, és hatékony lenne-e?

A büntetőjogi büntetés leggyakrabban megfogalmazott indokai a megtorlás, a cselekvőképtelenség megszüntetése, az elrettentés és a rehabilitáció.⁵⁸ A megtorlás a büntetés legrégebbi oka, amely az áldozat bosszúvágyának társadalmi demonstrációjává alakítja, hogy a rossznak következményei vannak.⁵⁹ Ezeknek a következményeknek a mérése a *lex talionisban* volt a legszorosabb: szemet szemért, fogat fogért. Egy vállalat - vagy egy elektronikus "személy" - megbírságolásának demonstratív hatása előnyösebb lehet, mint egy egyébként büntetlenül maradó bűncselekmény.⁶⁰

A büntetés-végrehajtási rendszer arra is felhasználható, hogy a bűncselekményért elítélteket cselekvőképtelenné tegye, fizikailag megakadályozva őket a bűnismétlésben. Ez általában a bebörtönzés különböző formáin keresztül történik, de magában foglalhatja a száműzetést, a végtagok amputálását, a kasztrálást és a kivégzést is. Társaságok esetében a működési engedély visszavonása vagy a kényszertörlesztés elrendelése is előfordulhat.⁶¹ Itt a veszélyes állatokkal és gépekkel való bánásmóddal való közvetlen analógiák is lehetnek

⁵⁷ A vállalatok számára ez néha nehéz, de nem leküzdhetetlen kihívásnak bizonyult. Lásd VS Khanna, "Corporate Criminal Liability: What Purpose Does It Serve? (1996) 109 Harvard L Rev 1513; JW Yockey, "Beyond Yates: From Engagement to Accountability in Corporate Crime" (2016) 12 New York U J of L and Business 412- 413; SW Buell, "Criminally Bad Management" in J Arlen (szerk.), *Research Handbook on Corporate Crime and Financial Misdealing* (Edward Elgar 2018) 59. Az a tény, hogy a vállalatok a szabad akarat vagy az erkölcsi felelősség hiánya ellenére is képesek büntetőjogi törvények megsértésére, feltehetően elveti ezt mint az AI-rendszerek büntetőjogi felelőssége elleni érvet.

⁵⁸ Vitatható, hogy a büntetőjog szimbolikus szerepe nem feltétlenül igényel tényleges büntetést - nem ritka, hogy a gyakorlatban nem hajtják végre a törvényeket. Ez azonban jellemzően arra a kifejezett vagy hallgatóságos döntésre támaszkodik, hogy bizonyos bűncselekményeket nem vizsgálnak ki vagy nem indítanak büntetőeljárást, nem pedig annak elfogadására, hogy a szereplők egy csoportját egyáltalán nem lehet büntetni.

⁵⁹ A feljelentést néha úgy mutatják be, mint a büntetés önálló indoklását. Lásd B Wringer, *An Expressive Theory of Punishment* (Palgrave Macmillan 2016).

⁶⁰ C Mulligan, "Revenge Against Robots" (2018) South Carolina L Rev 579; Hu (18. o.) 503-07.

⁶¹ WR Thomas, "Incapacitating Criminal Corporations" (2019) Vanderbilt L Rev. 905.

bár az olyan intézkedések, mint például egy rosszindulatú kutya elaltatása vagy egy hibás jármű forgalomból való kivonása inkább adminisztratív, mint büntetőjogi jellegűek, és nem függenek a "bűnösség" megállapításától.⁶² Egyes jogrendszerekben a gyermekek és a mentálisan sérült személyek bűncselekmény elkövetésére képtelennek minősülnek, mégis az állam őrizetbe veheti őket, ha úgy ítélik meg, hogy veszélyt jelentenek önmagukra vagy a közösségre.⁶³ Az ilyen személyek nem veszítik el személyiségüket; a mesterséges intelligencia rendszerek esetében nem szükséges személyiséggel felruházni őket ahhoz, hogy elzáráshoz hasonló intézkedéseket lehessen alkalmazni, ha egy termék visszahívható vagy egy engedély visszavonható.

Az elrettentés a büntetés újabb indoklása, amely az elkövetők racionalitásán alapul. A büntetések kialakításával olyan költségeket ró a viselkedésre, amelyeknek a célja, hogy ellensúlyozzák a lehetséges előnyöket. A bűnözés gazdasági elemzésre való visszavezetésének képessége különösen alkalmazhatónak tűnik mind a vállalatokra, mind a mesterséges intelligencia rendszerekre. Az előbbiek esetében azonban az ösztönzők valójában az emberi vezetők ellen irányulnak, akik egyébként a vállalaton keresztül a személyes és a vállalati haszon érdekében is összehangoltan cselekednének.⁶⁴ Egy mesterséges intelligencia rendszer esetében a bírság elrettentő hatása csak akkor alakítaná a viselkedést, ha a rendszer programozása a gazdasági haszon maximalizálására törekedne, tekintet nélkül magára az alapul szolgáló büntetőjogra.

A büntetés utolsó indoklása a rehabilitáció. Az elrettentéshez és az elrettentéshez hasonlóan ez is előremutató, és célja a visszaesés csökkentése. Az elrettentéstől eltérően azonban inkább a bűncselekmény elkövetésére vonatkozó döntést, mintsem az elkövetési képességet kívánja befolyásolni; az elrettentéstől ⁶⁵eltérően pedig ez a befolyásolás inkább belső, mintsem külsőleg kíván hatni. ⁶⁶ A természetes személyekre vonatkozó rehabilitáció gyakran inkább elméletben, mint a gyakorlatban elfogadottnak tűnik; különösen az Egyesült Államokban az 1970-es években kiesett a népszerűségéből.⁶⁷ A vállalatok esetében azonban a befolyás egyértelműbb eszközei ösztönözték a szűkre szabott büntetésekkel való kísérletezést, amelyek célja a jó magatartás ösztönzése és a rossz magatartás visszaszorítása.⁶⁸ Egy ilyen megközelítés jól alkalmazhatónak tűnhet a mesterséges intelligencia rendszerek esetében, mivel a büntetőjog megsértése

⁶² Lásd pl. D Legge és S Brooman, *Law Relating to Animals* (Cavendish Publishing 2000).

⁶³ A Loughnan, *Manifeszt örület: Mental Incapacity in the Criminal Law* (OUP 2012).

⁶⁴ A Hamdani és A Klement, "Corporate Crime and Deterrence" (2008) 61 Stanford L Rev 271. Vö. RA Guttman, "Effective Compliance Means Imposing Individual Liability" (2018) Emory5 Corporate Governance and Accountability Rev. 77.

⁶⁵ J Bentham, "Panopticon Versus New South Wales" in J Bowring (szerk.), *The Works of Jeremy Bentham* (William Tait 1843) 4. kötet, 174.

Személyiség v43a (28-Aug-20)

⁶⁶ Lásd T Ward és S Maruna, *Rehabilitation* (Routledge 2007).

⁶⁷ Lásd pl. AW Alschuler, "The Changing Purposes of Criminal Punishment: A Retrospective on the Past Century and Some Thoughts About the Next" (2003) 70 U Chicago L Rev 9. Vö. FT Cullen és KE Gilbert, *Reaffirming Rehabilitation* (2. kiadás, Anderson 2013).

⁶⁸ ME Diamantis, "Óraműves vállalatok: A Character Theory of Corporate Punishment" (2018) Iowa103 L Rev. 507.

a hibák hibakeresése, nem pedig a bűnök megbüntetése.⁶⁹ A rehabilitáció nevelési aspektusát valóban közvetlenül a gépi tanulással analizálták egy, a témáról szóló, könyvhosszú értekezésben.⁷⁰ Mégis, sem jogi személyiségre, sem az állam kényszerítő hatalmára nem lenne szükség ahhoz, hogy a gépi tanulás olyan kimenetekhez vezessen, amelyek nem sértik a büntetőjogot.

C. *Nincs lélek, hogy elkárhozzon*

Míg a vállalatok felelősségét igazoló érvek általában instrumentálisak, feltűnő, hogy a "robotbűnözőkről" szóló újonnan megjelenő szakirodalom hogyan csúszik át az antropomorfizmusba. Már maga a kifejezés is azt sugallja, hogy a humanoid mesterséges intelligencia rendszereket magasabb mércével kívánják mérni, mint például a különböző mértékben autonóm háztartási gépeket vagy a felhőben működő, testetlen mesterséges intelligencia rendszereket.⁷¹ Az ilyen megkülönböztetésnek nincs elvi oka, de ez a mesterséges intelligencia személyiséggel kapcsolatos, az instrumentális és az eredendő indoklást vegyítő érvekben rejlő feszültségre utal.

Érdekes módon a vállalatok jogi személyiségéről szóló viták általában az ellenkező problémára összpontosítanak: az emberektől való *különbözőségükre*, amelyet az első Thurlow báró frappánsan úgy jellemezett, hogy "nincs lelkük, hogy elkárhozzanak, és nincs testük, hogy megrúgják őket".⁷² A lélek hiánya nem akadályozta a vállalatok jogi személyiségét, és nem jelent elvi akadályt a mesterséges intelligencia rendszerek hasonló kezelésének. A vállalati személyiség azonban abban különbözik az AI személyiségétől, hogy a vállalatot emberek alkotják, akik révén működik, míg az *AI-rendszereket* emberek alkotják.⁷³

Az instrumentális okok tehát indokolhatják, hogy a mesterséges intelligencia rendszereknek jogi személyiséget tulajdonítsanak. De nem igénylik ezt. Az implicit antropomorfizmus elkerüli a további kihívásokat, például a személyiség küszöbének meghatározását, amikor az AI-rendszerek egy spektrumon belül léteznek, valamint azt, hogy a személyiség hogyan alkalmazható az elosztott rendszerekre. Lehetséges lenne tehát a vállalatokhoz hasonló jogi személyeket létrehozni - minden egyes autonóm jármű, intelligens orvosi eszköz, önletrajz-ellenőrző algoritmus és így tovább be lehetne építeni.⁷⁴ Ha valóban vannak felelősségi hiányosságok

⁶⁹ MA Lemley és B Casey, "Remedies for Robots" (2019) U86 Chicago L Rev 1370.

⁷⁰ Hallevey (18. o.) 210-11.

⁷¹ Vö. JM Balkin, "The Three Laws of Robotics in the Age of Big Data" (A robotika három törvénye a nagy adatok korában) (2017) Ohio78 State LJ 1219.

⁷² MA King, *Public Policy and the Corporation* (Chapman and Hall 1977) 1. Lásd pl. JC Coffee, Jr, "'No Soul to

Damn: No Body to Kick": An Unscandalized Inquiry into the Problem of Corporate Punishment (1981) 79 Michigan L Rev. 386.

⁷³ SM Solaiman, "A robotok, vállalatok, bálványok és csimpánzok jogi személyisége: A Quest for Legitimacy" (2017) 25 Artificial Intelligence and Law 174.

⁷⁴ Lásd pl. Chopra és White (41. o.). 161.

akkor lehetséges, hogy ilyen jogi formák tölthetnék ki őket. Egy ilyen megállapodásnak azonban inkább a termelők és a felhasználók lennének a haszonélvezői, akik így mentesülnének a felelősség egy része vagy egésze alól.

II. TERMÉSZETES SZEMÉLYISÉG: COGITO, ERGO SUM?

Az instrumentalizmus azonban nem az egyetlen ok, amiért a jogrendszerek elismerik a személyiséget. A természetes személyek esetében nem kell Turing-tesztet teljesíteni: a pusztán tény, hogy valaki megszületik, feljogosítja a személyiségre a törvény előtt.⁷⁵

Ez nem volt mindig így. Az emberi történelem nagy részében a rabszolgákat tulajdonként vették és adták el; az⁷⁶ őslakos népeket a földön kóborló állatokhoz hasonlították, ami igazolta a kisajátításukat;⁷⁷ és az angol jogban évszázadokon át Blackstone összefoglalója a nők helyzetéről úgy szólt, hogy "a férj és a feleség egy személy, és a férj az a személy".⁷⁸ A természetes személyek még ma is csak akkor élveznek teljes körű jogokat és kötelezettségeket, ha nagykorúak, épelméjűek, és nincsenek bebörtönözve.

Amint azt már korábban jeleztük, a mesterséges intelligencia személyisége mellett szóló érvek közül sokan hallgatólagosan vagy kifejezetten azt feltételezik, hogy a mesterséges intelligencia rendszerek olyan módon közelítik meg az emberi tulajdonságokat, amely feljogosítja őket a jog előtti hasonló elismerésre. Az ilyen érveket mind elemzésük, mind következményeik tekintetében megkérdőjelezték. Az elemzést illetően Neil Richards és William Smart "android tévhitnek" nevezte az AI-rendszerek antropomorfizálására irányuló tendenciát.⁷⁹ Kísérletről kísérletre kimutatták, hogy az emberek nagyobb valószínűséggel tulajdonítanak emberi tulajdonságokat - például erkölcsi érzékenységet - a gépeknek humanoid megjelenésük, természetes nyelvi kommunikációjuk, vagy pusztán a tény alapján, hogy nevet kaptak.⁸⁰ Bővebben

⁷⁵ Lásd pl. az Emberi Jogok Egyetemes Nyilatkozata, GA Res 217A(III) (1948), UN Doc A/810 (1948), 1. cikk; Polgári és Politikai Jogok Nemzetközi Egyezségokmánya (ICCPR), 16UNTS1966,999 197623. március,171, 6. cikk (1) bekezdés.

⁷⁶ Lásd J. Allain (szerk.), *A rabszolgaság jogi értelmezése: From the Historical to the Contemporary* (OUP 2013).

⁷⁷ S Chesterman, ""Csontváz jogi elvek": (1998) J40 Legal Pluralism and Unofficial L 61.

⁷⁸ L Holcombe, *Feleségek és tulajdon: századi Angliában* (Martin Robertson 1983) 18.

⁷⁹ NM Richards és WD Smart, "How Should the Law Think About Robots?" in R Calo, AM Froomkin, and I Kerr (szerk.), *Robot Law* (Edward Elgar 2016) 18-21.

⁸⁰ Vö. L Damiano és P Dumouchel, "Anthropomorphism in Human-Robot Co-evolution" (2018) 9 *Frontiers in Psychology* 468.

az emberi tulajdonságokat megközelítő mesterséges intelligenciával kapcsolatos komoly érvek megkérdőjeleződnek, mivel nem vizsgált feltételezések vannak arra vonatkozóan, hogy ezek a tulajdonságok hogyan nyilvánulnak meg az emberekben.⁸¹

Ami a következményeket illeti, az Európai 2017Parlament állásfoglalása arra készítette a kontinens több száz AI-szakértőjét, hogy nyílt levélben figyelmeztessenek arra, hogy a mesterséges intelligencia jogi személyisége "etikai és jogi szempontból" nem lenne megfelelő. Érdekes módon az ilyen figyelmeztetések maguk is beleeshetnek az androidos tévedésbe, mivel azt feltételezik, hogy a természetes személy modelljén alapuló jogi státusz szükségszerűen magával hozza az uniós jog által garantált "emberi" jogokat.⁸² Más szerzők őszintén elismerik, hogy a mesterséges intelligenciával működő rendszerek személyiségének megtagadásának egyetlen alapja végső soron egyfajta fajgyűlölet lehet - az emberi jólét előnyben részesítése a robotok jólétével szemben, mivel mi, a jogalkotók is emberek vagyunk.⁸³ Ha a mesterséges intelligencia rendszerek olyan kifinomulttá válnak, hogy ez lesz a legerősebb védekezésünk, akkor a probléma nem a jogi státuszukkal, hanem a sajátunkkal lesz.⁸⁴

Ez a szakasz mindazonáltal komolyan veszi azt a gondolatot, hogy bizonyos mesterséges intelligencia-rendszerek eredendő tulajdonságaik miatt jogosultak lehetnek a személyiségre. Annak technikai aspektusai, hogy ezek a tulajdonságok hogyan nyilvánulhatnak meg - és az általuk utánczott emberi tulajdonságok részletes vizsgálata - meghaladja e cikk kereteit.⁸⁵ Ehelyett arra fogunk összpontosítani, hogy hogyan és miért lehetne kiterjeszteni a természetes személyiséget. Az első megvizsgálandó kérdés az, hogyan kezelték ezt a múltban, például a sokáig a fehérekkel szemben alacsonyabb rendűnek tekintett természetes személyek jogainak és jogosítványainak megadásával. A közelmúltban aktivisták és tudósok szorgalmazták bizonyos jogok további kiterjesztését a nem emberi állatokra, például a csimpánzokra, saját eredendő tulajdonságaik alapján. A vizsgálat ezután rátér a mesterséges intelligenciával működő rendszerek számára a ma legerőteljesebben megfogalmazott, nem instrumentális, hanem eredendő okokból biztosított jogokra: arra, hogy a mesterséges intelligenciával működő rendszerek képesek legyenek alkotásaik tulajdonjogát megszerezni.

⁸¹ Lásd pl. E Hildt, "Mesterséges intelligencia: Does Consciousness Matter? (2019) 10(1535) Frontiers in Psychology, 1-3; G Meissner, 'Artificial Intelligence: Consciousness and Conscience' (2020) 35 AI & Society 231. John Searle híres "kínai szoba" érveléséhez lásd JR Searle, "Minds, Brains, and Programs" (1980) 3 Behavioral

and Brain Sciences 417-24.

⁸² Nyílt levél az Európai Bizottságnak: Mesterséges intelligencia és robotika (2018. április), (2) bekezdés b) pont. Vö. Turner (40. o.) 189-90. o.

⁸³ Lásd pl. Bryson, Diamantis és Grant (27. o.) 283. o. Vö. P Singer, "Speciesism and Moral Status" (2009) 40 *Metaphilosophy* 567.

⁸⁴ Lásd alább a III.

⁸⁵ Lásd pl. J-M Fellous és MA Arbib, *Who Needs Emotions? The Brain Meets the Robot* (OUP 2005); W Wallach és C Allen, *Moral Machines: Teaching Robots Right from Wrong* (OUP 2009).

A. *A természetes személyiség kiterjesztése*

Az erkölcsi világegyetem íve hosszú, ahogyan Dr. Martin Luther King Jr. híres intonációja szerint, de az igazságosság felé hajlik. Amikor az Egyesült Államok 1776-ban megfogalmazta Függetlenségi Nyilatkozatát, az a felfogás, hogy "minden ember" [sic] "egyenlőnek teremtett", bizonyíthatóan nem volt igaz. Egy évtizeddel később a francia Emberi Jogok Nyilatkozata hasonlóképpen természetes és megmásíthatatlan jogokat hirdetett mindenki számára - Jeremy Bentham megjegyzése szerint ez "gólyalábakon álló képtelenség".⁸⁶ Az ember ugyan szabadnak születhetett, ahogyan Rousseau vélekedett *A társadalmi szerződés* első soraiban, de mindenütt láncra verve maradt.⁸⁷

A következő évszázadokban azonban fokozatosan megvalósultak ezek a magasztos törekvések, és a jogok elterjedtek. A huszadik század közepére az Emberi Jogok Egyetemes Nyilatkozata már azt állíthatta, hogy minden ember "szabadnak, méltóságában és jogaiban egyenlőnek született", annak ellenére, hogy egyharmaduk olyan területeken élt, amelyeket maga az ENSZ is az önrendelkezési joggal nem rendelkezőnek minősített. Ezt követte a dekolonizáció, az apartheid felszámolása, a nők felszabadítása és más mozgalmak; a jogok továbbra is vitatottak, de ma már gyakorlatilag egyetlen állam sem állítaná komolyan, hogy az emberi felnőttek nem személyek a törvény előtt.⁸⁸

Érdekes módon a mesterséges intelligencia jogi személyisége mellett szóló érvek egy része nem a természetes személyiség progresszív narratívájára, hanem a rabszolgaság sötét történetére támaszkodik. Andrew Katz és Ugo Pagallo például analógiát talál az ókori római jog *peculium* mechanizmusával, amely szerint a rabszolga nem rendelkezett jogi személyiséggel, mégis több lehetett, mint gazdája egyszerű ügynöke.⁸⁹ A személyiség kreatív értelmezésének példajaként ez érdekes, bár inkább instrumentális indokokra támaszkodik, mint a rabszolgák eredendő tulajdonságaira. Amint Pagallo megjegyzi, *a peculium* valójában egyfajta "korlátolt felelősségű társaság prototípusa" volt.⁹⁰ Az előző szakaszból kiderül, hogy ma már nincs akadálya annak, hogy a jogrendszerek ilyen struktúrákat hozzanak létre -

⁸⁶ J Bentham, "Anarchical Fallacies" in J Bowring (szerk.), *The Works of Jeremy Bentham* (William Tait 1843) 2. kötet, 501.

⁸⁷ J-J Rousseau, *A társadalmi szerződés* (GDH Cole tr, először megjelent JM1762, Dent 1923) 49.

⁸⁸ A hitehagyottakra és a fogyatékkal élőkre vonatkozó korlátozott kivételekről lásd PM Taylor, *A Commentary on the International Covenant on Civil and Political Rights* (CUP 2020) 449-54. Az anekdotalis csecsemőkkel kapcsolatos vitát lásd Kurki (15. sz.). 9.

⁸⁹ A Katz, "Intelligent Agents and Internet Commerce in Ancient Rome" (2008) *Society 20for Computers and Law* 35; U Pagallo, *The Laws of Robots: Crimes, Contracts, and Torts* (Springer 2013) 103-06. Lásd még H Ashrafian, "Artificial Intelligence and Robot Responsibilities: Innovating Beyond Rights" (2015) 21 *Science and Engineering Ethics* 325; S Nasarre-Aznar, "Ownership at Stake (Once Again): Housing, Digital Contents, Animals, Személyiség v43a (28-Aug-20)

and Robots" (2018) 10 J Property, Planning, and Environmental L 78; E Fosch-Villaronga, *Robots, Healthcare, and the Law: Regulating Automation in Personal Care* (Routledge 2019) 152. 2017-ben Franciaországban Peculium néven digitális bankot hoztak létre - feltehetően olyan befektetők számára, akik soha nem tanultak latinul.

⁹⁰ Pagallo (n 89) 104.

Ami azt illeti, hogy ezt *kell-e* tenniük, a rabszolgasághoz kapcsolódó, régen elvetett törvényekre való hivatkozás nem biztos, hogy a lehető legerősebb érv.⁹¹

Egy alternatív megközelítés az, hogy megvizsgáljuk, hogyan bánik a jogrendszer az állatokkal.⁹² Többnyire úgy tekintenek rájuk, mint megvásárolható és eladható vagyontárgyakra, de úgy is, mint amelyek megérdemlik az "emberséges" bánásmódot.⁹³ Az állatok által okozott károkért a tulajdonosok felelőssége csak korlátozottan alkalmazható a mesterséges intelligencia rendszerekre;⁹⁴ Itt a kérdés az, hogy ezek az állatok "birtokolhatják-e" önmagukat.

Számos kísérletet tettek arra, hogy a nem emberi állatoknak személyiséget tulajdonítsanak, kevés sikerrel. 2013-ban például a Nonhuman Rights Project négy fogságban tartott csimpánz nevében indított pert, azzal érvelve, hogy az állatok fejlett kognitív képességekkel, autonómiával és öntudattal rendelkeznek. A New York Állami Bíróság Fellebbviteli Osztálya a habeas corpus iránti kérelmeket elutasítva nem vitatta ezeket a tulajdonságokat, de úgy ítélte meg, hogy a személyiséghez hasonló jogok kiterjesztése hagyományosan a kötelezettségek társadalmi szerződés formájában történő előírásához kapcsolódik. Mivel "mondanom sem kell", hogy a csimpánzok nem viselhetek jogi kötelezettségeket, nem élvezhettek olyan személyiségi jogokat, mint a szabadsághoz való jog.⁹⁵ Ez furcsa alap volt az ügy elutasításához, mivel sok olyan ember, aki nem képes jogokat vagy kötelezettségeket gyakorolni - csecsemők, kómában fekvő személyek -, mégis személynek minősül a törvény előtt.⁹⁶ Egy párhuzamos ügyben ezt az érvelést azzal a körkörös indokkal utasították el, hogy "figyelmen kívül hagyja azt a tényt, hogy ezek még mindig emberi lények, az emberi közösség tagjai".⁹⁷ Az állami fellebbviteli bírósághoz való fellebbezés engedélyezését elutasították, de az egyik bíró egyetértő véleményt adott ki, amely az ilyen perek jövőjéről szóló spekulatív megjegyzéssel zárult. A kérdés, Fahey bíró

⁹¹ M Chinen, *Jog és autonóm gépek: The Co-Evolution of Legal Responsibility and Technology* (Edward Elgar 2019) 19.

⁹² Lásd pl. VAJ Kurki és T Pietrzykowski (szerk.), *Jogi személyiség: Animals, Artificial Intelligence and the Unborn* (Springer 2017); S Stucki, "Towards a Theory of Legal Animal Rights: Simple and Fundamental Rights" (2020) OJLS megjelenés alatt.

⁹³ Vö. K Sykes, "Emberi dráma, állati kísérletek: What the Medieval Animal Trials Can Teach Us About Justice for Animals" (2011) 17 *Animal Law* 273.

⁹⁴ A "veszélyes fajhoz" tartozó állat esetében az állattartóról feltételezhető, hogy tudja, hogy az állat hajlamos kárt okozni, és az általa okozott kárért felelősségre vonható, anélkül, hogy az állattartó hibáját bizonyítani kellene. Más állatok esetében bizonyítani kell, hogy az állattartó tudta, hogy az adott állat veszélyes. Ezeket az állatokra vonatkozó szokásjogi szabályokat ma már jogszabályok szabályozzák: Állatokról szóló 1971. évi törvény (Egyesült Királyság). Az angol bíróságok elzárkóztak a szigorú felelősség általános doktrínájától a "rendkívül veszélyes tevékenységek" esetében. Lásd: C Witting, *Street on Torts* (15th edn, OUP 2018) 453.

⁹⁵ *People ex rel Nonhuman Rights Project, Inc. kontra Lavery*, NYS998 2d (248App Div, 2014).

⁹⁶ RS Abate, *Az éghajlatváltozás és a hangtalanok: A jövő nemzedékek, a vadon élő állatok és a természeti erőforrások védelme*

(CUP 2019) 101-02; Kurki (15. o.) 6-10 (megkülönböztetve a passzív és az aktív jogi személyiséget).

⁹⁷ *Nonhuman Rights Project, Inc ex rel Tommy v Lavery*, NYS54 3d (392,App 396Div, 2017).

megfigyelt, mélyreható és messzemenő. "Végső soron nem fogjuk tudni figyelmen kívül hagyni. Bár vitatható, hogy a csimpánz nem "személy", de kétségtelen, hogy nem csupán egy dolog".⁹⁸

Gabriel Hallevy amellett érvelt, hogy az állatok közelebb állnak az emberhez, mint a mesterséges intelligencia rendszerek, ha az érzelmeket vesszük figyelembe a racionalitással szemben, de ez általában nem vezetett ahhoz, hogy a jog szerint személyiséget kapjanak. Ehelyett az AI-rendszerek *racionalitása* az, ami a személyiség alapját képezi.⁹⁹ Ez igaz lehet a bűncselekmény mentális elemének megállapítására való képesség tekintetében, de az a tény, hogy egy csimpánz kínzása bűncselekménynek minősül, de egy számítógépé nem, szintén fontos különbségre utal abban, ahogyan a jogrendszer a kétféle entitást értékeli. Valójában erősebb érv szólhat a megtestesült mesterséges intelligencia rendszerek védelme mellett, amelyek érzelmi reakciókat váltanak ki az emberekből - függetlenül belső feldolgozásuk kifinomultságától. Valószínűnek tűnik, hogy az ilyen "szociális robotok" védelmére valamikor törvényeket fognak elfogadni, hasonlóan az állatokkal szembeni bántalmazásról szóló törvényekhez. Az említett törvényekhez hasonlóan a védelmet valószínűleg inkább a társadalmi erkölcsök fogják vezérelni, mint a következetes biológiai - vagy technológiai - normák.¹⁰⁰

Az a feltételezés, hogy a természetes jogalanyiség az emberekre korlátozódik, olyan mélyen gyökerezik a legtöbb jogrendszerben, hogy még csak ki sem fogalmazódik.¹⁰¹ Az, hogy még a legközelebbi evolúciós rokonainkra sem terjesztenek ki hasonló jogokat, rosszat sejtet a feltételezett eredendő tulajdonságokon alapuló mesterséges intelligencia-személyiség hívei számára.¹⁰²

B. *A kreativitás jutalmazása*

A mesterséges intelligencia rendszerek személyként való elismerésének megfontolásának egyik oka nem arra összpontosít, hogy mik azok, hanem arra, hogy mire képesek. Ez többnyire úgy fogalmazódik meg, mint az a kérdés, hogy egy magánszemély vagy egy vállalat igényelhet-e tulajdonjogot az AI-rendszer által végzett munkára. Az ilyen vitákban azonban hallgatólagosan vagy kifejezetten az a felfogás jelenik meg, hogy ha az ilyen munkát egy ember végezte volna, akkor az ember maga is tulajdonosa lenne a munkának.

⁹⁸ *Nonhuman Rights Project, Inc ex rel Tommy v Lavery*, NE 1003d (846,NY848, 2018).

⁹⁹ Hallevy (n18) 28.

¹⁰⁰ K Darling, "A jogi védelem kiterjesztése a szociális robotokra: The Effects of Anthropomorphism, Empathy,

and Violent Behavior Towards Robotic Objects" in R Calo, AM Froomkin, and I Kerr (eds), *Robot Law* (Edward Elgar 2016) 226-29. Vö. R Calo, "Robotics and the Lessons of Cyberlaw" (2015) *California* 103 L Rev. 532.

¹⁰¹ Kurki (n 15) 8.

¹⁰² Vö. *ibid.* 176-78. (annak megvitatása, hogy a mesterséges intelligencia lehet-e "végső soron értékes", és így jogosult-e a személyiségre).

Valójában már régóta kérdéses, hogy a gépi segítséggel történő alkotás védhető-e szerzői jogi védelemmel.¹⁰³ A korai fényképeket például nem védték, mert a fény puszta megörökítése egy *camera obscura* lencséjén keresztül nem számított valódi szerzői alkotásnak.¹⁰⁴ Oscar Wilde ikonikus képének egészen az Egyesült Államok Legfelsőbb Bíróságáig kellett eljutnia ahhoz, hogy a gépi úton előállított alkotások szerzői jogát elismerjék.¹⁰⁵ Ma már más a kihívás: nem az a kérdés, hogy a fényképész "birtokolhatja-e" a gép által passzívan rögzített képet, hanem az, hogy ki lehet a gép által aktívan létrehozott új alkotások tulajdonosa. Egy számítógépes program, mint például egy szövegszerkesztő, nem tulajdonolja a rajta begépelte szöveget, mint ahogyan egy toll sem az általa írt szavak tulajdonosa. A mesterséges intelligenciával működő rendszerek azonban ma már híradásokat írnak, dalokat komponálnak, képeket festenek - ezek a tevékenységek értéket teremtenek, de vajon élvezhetik-e és kell-e rájuk a szerzői jog védelmét?

A legtöbb joghatóságban a válasz nem.

Az Egyesült Államok Szerzői Jogi Hivatala például kijelentette, hogy az "eredeti szerzői művek"¹⁰⁶ jogi védelme az "ember által létrehozott" művekre korlátozódik. A hivatal nem veszi nyilvántartásba azokat a műveket, amelyeket "gép vagy pusztán mechanikus folyamat állít elő, amely véletlenszerűen vagy automatikusan működik, emberi szerző alkotó közreműködése vagy beavatkozása nélkül".¹⁰⁷ A "bármilyen" szó kulcsfontosságú, és felveti a kérdést, hogy milyen szintű emberi közreműködés szükséges a szerzőség megállapításához.¹⁰⁸

Gondoljunk csak a világ leghíresebb szelfijére - egy fekete címeres makákóról. David Slater Indonéziába utazott, hogy lefotózza a veszélyeztetett majmokat, amelyek túl idegesek voltak ahhoz, hogy közelképeket készítsen róluk. Ezért felállított egy olyan kamerát, amely lehetővé tette számukra, hogy saját maguk készítsenek fotókat.¹⁰⁹ Miután a képek jelentős nyilvánosságot kaptak, az állatvédők azzal érveltek, hogy a majmok nagyobb joggal tartanak igényt a fényképek szerzői jogára, mint a kamera tulajdonosa. Slater

¹⁰³ J Grimmelmann, "There's No Such Thing as a Computer-Authored Work - and It's a Good Thing, Too" (2016) Columbia 39J L & the Arts 403.

¹⁰⁴ M de Cock Buning, "Mesterséges intelligencia és a kreatív ipar: New Challenges for the EU Paradigm for Art and Technology by Autonomous Creation" in W Barfield and U Pagallo (szerk.), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar 2018) 524.

¹⁰⁵ *Burrow-Giles Lithographic Co v Sarony*, 111 US 53 (1884). A viták azonban folytatódtak, és Németország 1965-ig visszatartotta a fényképek teljes szerzői jogát. A Nordemann, "Germany" in Y Gendreau, A Nordemann, and R Oesch (szerk.), *Copyright and Photographs. An International Survey* (Kluwer 1999) 135.

¹⁰⁶ 17 USC §102(a).

¹⁰⁷ *Compendium of US Copyright Office Practices* (3. kiadás, US Copyright Office 2019), 313.2. § (kiemelés hozzáadva).

¹⁰⁸ Lásd DJ Gervais, "The Machine as Author" (2020) 105 Iowa L Rev forthcoming (az "eredetiség okozati összefüggésének" tesztjét javasolja).

¹⁰⁹ C Cheesman, 'Photographer Goes Ape Goes over Monkey Selfie: Kié a szerzői jog?', *Amateur Photographer*, 2014. augusztus 7.

végül a hatályos törvények alapján nyert¹¹⁰- bár a megegyezés részeként beleegyezett, hogy a képekből származó jövőbeli jogdíjak százalékát 25 a címeres makákókat védő csoportoknak adományozza.¹¹¹ Ahogy a számítógépek egyre több tartalmat generálnak az emberi programozóktól függetlenül, az embereknek egyre nehezebb lesz learatniuk a babérokat. Ahelyett, hogy egy majmot betanítanánk arra, hogyan kell megnyomni egy gombot, ez inkább olyan lehet, mintha egy tanár próbálná magáénak tudni a tanítványa munkáját.

Ami azt a normatív kérdést illeti, hogy maguknak a mesterséges intelligencia *rendszernek is legyen-e* igényük a tulajdonjogra, a szerzői jog mögött álló politikát gyakran úgy fogalmazzák meg, mint az innováció ösztönzését. Ezt régóta sürgősnek vagy nem megfelelőnek tartják a számítógépek esetében. Pamela Samuelson írta: "Mindössze elektromosságra (vagy más mozgatóerőre 1986,) van szükség ahhoz, hogy a gépek termelésbe álljanak".¹¹² A Turing-teszt itt egy másfajta gondolkísérletet kínál: minél inkább úgy tervezik a gépeket, hogy emberi tulajdonságokat másoljanak, annál fontosabbá válhatnak az ilyen ösztönzők.

A közelmúltig Kína azt az ortodox nézetet követte, hogy a mesterséges intelligenciával előállított művek nem jogosultak szerzői jogi védelemre.¹¹³ 2019 decemberében azonban egy kínai kerületi bíróság kimondta, hogy egy algoritmus által előállított cikket nem lehet engedély nélkül másolni. A szóban forgó cikk egy pénzügyi jelentés volt, amelyet a Tencent tett közzé azzal a megjegyzéssel, hogy azt a Dreamwriter, a vállalat által kifejlesztett hírszerkesztő program "automatikusan írta" Sanghajban. A 2015. Yingxun Technology Company engedély nélkül lemásolta a cikket, és a Tencent beperelte. A cikket levették, de a jogsértő céget 1500 jen (216 dollár) megfizetésére kötelezték "gazdasági veszteség és jogvédelem" címén.¹¹⁴

A kínai eset a szerzői jog elismerésének egy külön okát tükrözi, amely a kreatív folyamatokba történő előzetes befektetések védelme. Ez az indoklás azt feltételezi, hogy ilyen védelem hiányában a beruházások elapadnak, és csökken a kreatív alkotások kínálata.

¹¹⁰ *Naruto v Slater*, F 8883d (4189th Cir, 2018). A bíróság megállapította, hogy a Naruto nem rendelkezik az amerikai szerzői jogi törvény szerinti perlési joggal, és nem tart igényt a Slater által közzétett fényképekre.

¹¹¹ M Haag, 'Kié a majom szelfi? Settlement Should Leave Him Smiling', *New York Times* (2017. szeptember 11.).

¹¹² P Samuelson, "Allocating Ownership Rights in Computer-Generated Works" (1986) U 47Pittsburgh L Rev. 1199.

¹¹³ *Beijing Feilin Law Firm kontra Baidu Corporation (No. 239)* (2019. április 25.) (Beijing Internet Court); CHEN Ming, "Beijing Internet Court Denies Copyright to Works Created Solely by Artificial Intelligence" (2019) J14 Intellectual Property L & Practice. 593.

¹¹⁴ 深圳市腾讯计算机系统有限公司 [Shenzhen Tencent Computer System Co Ltd] kontra 上海盈讯科技有限公司 [Shanghai Yingxun Technology Co Ltd] (december 2019 24.) (Shenzhen Nanshan District People's Court); Személyiség v43a (28-Aug-20)

ZHANG Yangfei, "Court Rules AI-Written Article Has Copyright", *China Daily* (2020. január 9.).

működik.¹¹⁵ A szerzői jog ilyen megközelítése nagyjából összhangban van a munkavégzés során létrehozott művekre vonatkozó szokásjogi doktrínákkal, amelyek az Egyesült Államokban bérmunkaként ismertek, és amelyek szerint a szerzői jog a munkáltatót vagy a művet megrendelő magánszemélyt illeti meg annak ellenére, hogy a tényleges "szerző" valaki más.¹¹⁶ Ez nem feltétlenül áll rendelkezésre a polgári jogi jogrendszerekben, amelyek nagyobb hangsúlyt fektetnek az emberi szerző erkölcsi jogaira.¹¹⁷

Nagy-Britanniában az ott elfogadott jogszabályok 1988-ban szerzői jogi védelmet biztosítanak a "számítógéppel létrehozott" művek számára, amelyek "szerzőjének" azt a személyt tekintik, aki "a mű létrehozásához szükséges intézkedéseket" megtette.¹¹⁸ Hasonló jogszabályokat fogadtak el Új-Zélandon,¹¹⁹ Indiában,¹²⁰ Hongkongban¹²¹ és Írországban.¹²² Bár viták merülhetnek fel azzal kapcsolatban, hogy ki végezte a "szükséges intézkedéseket", az egyetlen lehetséges kimenetel az, hogy a tulajdonjog egy elismert jogi személyt vagy senkit sem illet meg.¹²³

Az Európai Parlament 2020 áprilisában jelentéstervezetet adott ki, amelyben amellet érvelt, hogy a mesterséges intelligencia által létrehozott művek a szellemi alkotásokkal "egyenértékűnek" tekinthetők, és ezért szerzői jogi védelem alatt állnak. Ellenezte azonban, hogy magának a mesterséges intelligenciának bármilyen személyiséget adjanak, és azt javasolta, hogy a tulajdonjog ehelyett "azt a személyt illeti meg, aki a művet jogszerűen elkészíti és közzéteszi, feltéve, hogy a technológia tervezője nem tartotta fenn kifejezetten a mű ilyen módon történő felhasználásának jogát".¹²⁴ Érdekes a szellemi alkotással való "egyenértékűség", amelyet itt az alkotói folyamat helyett az "alkotói eredményen" alapuló alkotások elismerésének javasolt elmozdulása indokol.¹²⁵

¹¹⁵ K Raustiala és CJ Sprigman, "The Second Digital Disruption: Streaming and the Dawn of Data-driven Creativity" (2019) 94 New York U L Rev 1603-04.

¹¹⁶ Vö. *Asia Pacific Publishing Pte Ltd kontra Pioneers & Leaders (Publishers) Pte Ltd* [2011] 4 SLR 381, 398-402 (Szingapúri Fellebbviteli Bíróság) (a szerzőség és a tulajdonjog megkülönböztetése).

¹¹⁷ D Lim, "AI & IP: Innovation & Creativity in an Age of Accelerated Change" (2018) Akron 52 L Rev 843-46.

¹¹⁸ Szerzői jogról, formatervezési mintákról és szabadalmakról szóló 1988. évi törvény (Egyesült Királyság), 9. szakasz (3) bekezdés. A 178. szakasz meghatározása szerint a "számítógéppel létrehozott" kifejezés azt jelenti, hogy a művet "számítógéppel hozták létre olyan körülmények között, hogy a műnek nincs emberi szerzője".

¹¹⁹ Szerzői jogi törvény (1994NZ), 5. szakasz (2) bekezdés a) pont.

¹²⁰ Szerzői jogról szóló törvény (1994India), szakasz 2.

¹²¹ Szerzői jogi rendelet (1997HK), 11. szakasz (3) bekezdés.

¹²² A szerzői és szomszédos jogokról szóló (2000Írországi) törvény 21. szakaszának f) pontja.

¹²³ Lásd pl. *Nova Productions kontra Mazooma Games* [2007] EWCA Civ 219; *A Brown et al, Contemporary*

Intellectual Property: Law and Policy (5. kiadás, OUP 2019) 100-01.

¹²⁴ S Séjourné, jelentéstervezet a szellemi tulajdonjogokról a mesterséges intelligencia technológiák fejlesztéséhez (Európai Parlament, Jogi Bizottság, 2020/2015(INI), 2020. április24), 9-10. bekezdés.

¹²⁵ Ibid., indokolás.

Egyelőre tehát a szerzői jog nem lehet a mesterséges intelligencia rendszerek tulajdonában - és nincs is rá szükség ahhoz, hogy elismerjék e rendszerek kreativitását. Mindazonáltal a "számítógéppel előállított" művekre vonatkozó korlátozott jogok nyilvánvalóan fenntartásokat tartalmaznak azzal kapcsolatban, hogy a tulajdonjogot bárki más követelheti. Az időtartam általában rövidebb, és a "szerzőnek" tekintett személy nem érvényesíthet erkölcsi jogokat - például a mű szerzőjeként való azonosításhoz való jogot.¹²⁶ A Szellemi Tulajdon Világszervezetének (WIPO) egyik kiadványa felismerte ezt a dilemmát, és megjegyezte, hogy az ilyen művek kizárása "az emberi kreativitás méltóságát a gépi kreativitással szemben előnyben részesítené", ami a fogyasztók számára minél több kreatív alkotás elérhetővé tételének rovására menne. A jelentés szerint a közéletet a "csökkentett védelmi idő és egyéb korlátozások" jelenthetnék.¹²⁷

C. *A feltalálók védelme*

Míg a szerzői jogban a vita arról folyik, hogy kié a mesterséges intelligencia rendszerek által létrehozott művek tulajdonjoga, addig a szabadalmi jogban az a kérdés, hogy egyáltalán lehet-e ezeket birtokolni. A szabadalmi jog a legtöbb joghatóságban előírja vagy feltételezi, hogy a "feltalálónak" embernek kell lennie. 2019 júliusában Stephen Thaler úgy döntött, hogy teszteli ezeket a feltételezéseket, és olyan szabadalmakat nyújtott be Nagy-Britanniában, az Európai Unióban és az Egyesült Államokban, amelyekben "feltalálóként" egy AI-rendszer, a DABUS szerepel.¹²⁸ A brit Szellemi Tulajdoni Hivatal hajlandó volt elfogadni, hogy a DABUS hozta létre a találmányokat, de a vonatkozó jogszabályok megkövetelték, hogy a feltaláló természetes személy legyen, ne pedig gép.¹²⁹ Az Európai Szabadalmi Hivatal (EPO) ugyanerre a célra egy sokkal körülményesebb utat követett, és elutasította a bejelentéseket azon az alapon, hogy egy gép feltalálónak való kijelölése nem felel meg a "formai követelményeknek". Ezek közé tartozott "a feltaláló családi és utónevének, valamint teljes lakcímének" feltüntetése.¹³⁰ Az EPO megjegyezte, hogy a név nem csak azonosítja a személyt: lehetővé teszi számára jogainak gyakorlását és formálja

¹²⁶ A szerzői jogról, formatervezési mintákról és szabadalmakról szóló törvény 12. szakaszának (7) bekezdése (az ilyen művek védelme a szerző halála 70 utáni évek helyett évekre 50 korlátozódik), 79. szakasza (kivételesen az erkölcsi jogok alól).

¹²⁷ Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence (Szellemi Tulajdon Világszervezet, WIPO/IP/AI/2/GE/20/1 REV, 2020. május 21.), 23. bekezdés. Lásd még: M du Sautoy, *A kreativitási kódex: Art and Innovation in the Age of AI* (Harvard UP 2019) 102; R Abbott, *The Reasonable Robot: Artificial Intelligence and the Law* (CUP 2020) 71-91.

¹²⁸ A szabadalmak egy "élelmiszer-tartályra" és "fokozott figyelemfelkeltő eszközökre és módszerekre" vonatkoztak. A DABUS a Device for the Autonomous Bootstrapping of Unified Sentience rövidítése.

¹²⁹ Az 1977. évi szabadalmi törvény (Egyesült Királyság) 7. és 13. szakasza. Lásd: *A feltaláló megnevezésére és a Személyiség v43a* (28-Aug-20)

szabadalmi bejelentéshez való jogra vonatkozó 7. és 13. szakasz követelményeinek teljesülése a GB1816909.4 és a GB1818161.0 tekintetében (BL O/741/19) (2019. december 4.) (UK Intellectual Property Office), 14-20. pont. A bíróság a továbbiakban megjegyezte, hogy a Thaler nem szerzhetett volna tulajdonjogot a DABUS-tól, "mivel a feltaláló maga nem lehet tulajdonosa" (23. bek.).

¹³⁰ Európai Szabadalmi Egyezmény, München, október 5 1973, hatályos október október 7 81, 19. cikk 1977, (1) bekezdés.

a személyiségük része. A "dolgoknak" ezzel szemben "nincsenek olyan jogaik, amelyeket egy név lehetővé tenné számukra, hogy gyakoroljanak."¹³¹

Az amerikai bejelentést szintén elutasították, részben arra a tényre alapozva, hogy a vonatkozó jogszabályok többször hivatkoztak a feltalálókra, akik "természetes személyekre jellemző névmásokat" használtak, mint például a "maga" és a "maga". A Szabadalmi és Kereskedelmi Hivatal (USPTO) olyan esetekre hivatkozott, amelyek szerint az elképzelés - "a feltalálói minőség próbaköve" - "mentális aktus", amely "a feltaláló elméjében" történik. Ezek az ügyek arra a következtetésre jutottak, hogy a feltalálás ebben az értelemben természetes személyekre korlátozódik, nem pedig társaságokra. Az USPTO arra a következtetésre jutott, hogy a mesterséges intelligenciát "feltalálóként" feltüntető bejelentés ezért hiányos, de óvatosan járt el azzal kapcsolatban, hogy ne állapítsa meg, "ki vagy mi" hozta létre ténylegesen a szóban forgó találmányokat.¹³²

Ezek a határozatok összhangban voltak az esetjoggal és a szabadalmi hivatalok gyakorlatával világszerte, amelyek közül - még - egyik sem teszi lehetővé, hogy egy mesterséges intelligencia rendszert feltalálóként ismerjenek el. A szerzői joggal analóg módon a szabadalmi rendszer egyik célja az innováció ösztönzése azáltal, hogy a nyilvánosságra hozatalért cserébe időben korlátozott monopóliumot biztosít. Amint azt még a DABUS alkotói is elismerték, egy mesterséges intelligencia rendszert nem valószínű, hogy a szabadalmi oltalom kilátásba helyezése motiválná az innovációra. Bármilyen ilyen motiváció a programozásában rejlik: a rendszert az innovációra kell utasítani.¹³³

Ami azt illeti, hogy egy emberi "feltalálónak" jóvá lehetne-e írni egy ilyen rendszer által végzett munkát, nincs megfelelője a bérmunka doktrínának. Ahhoz, hogy valaki feltaláló legyen, a találmányt ténylegesen az embernek kell kitalálnia.¹³⁴ Lehetségesek a közös találmányok, és a hozzájárulásoknak nem kell azonosnak lenniük, de ha nincs olyan természetes személy, aki jelentős koncepcionális hozzájárulást nyújt, akkor a találmány a jelenlegi jog szerint nem részesülhet szabadalmi oltalomban.¹³⁵

¹³¹ Az EPO EP bejelentéséről szóló határozatának indoklása (2020. 275163január 1827.) (Európai Szabadalmi Hivatal), (2) bekezdés.

22.

¹³² *In re Application No.: 16/524,350 (Petícióról szóló határozat)* (2020. április 22.) (US Patent and Trademark Office).

¹³³ *A feltaláló megnevezésére és a szabadalmi bejelentéshez való jogra vonatkozó 13. szakasz követelményeinek teljesülése a GB1816909.4 és a GB1818161.0 (129. sz.) tekintetében*, (1) bekezdés. 28.

¹³⁴ *Manual of Patent Examining Procedure (MPEP)* (9th edn, US Patent and Trademark Office 2017), §2137.01.

¹³⁵ JA Cubert és RGA Bone, "The Law of Intellectual Property Created by Artificial Intelligence" in W Barfield és

U Pagallo (szerk.), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar 2018) 418. Van egy gyenge érv, miszerint az értelmezési lehetőség abban rejlik, hogy a "feltalálót" az amerikai jogban úgy határozzák meg, mint azt a személyt, aki "feltalálja *vagy felfedezi*" a találmány tárgyát. 35 USC § 101. Lásd pl. R. Abbott, "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law" (2016) Boston 57 College L Rev 1098; WM Schuster, "Artificial Intelligence and Patent Ownership" (2018) Washington 75 & Lee L Rev 1977; K Hartung, "Dear USPTO: Patents for Inventions by AI Must Be Allowed", *IP Watchdog*, május. 212020.

A közelmúltbeli fejlemények érdekességei közé tartozik, hogy milyen eszközökkel jutottak ugyanerre a következtetésre. A szerzői joghoz hasonlóan a jelek szerint nem volt komoly kétség afelől, hogy egy mesterséges intelligencia rendszer képes olyan dolgokat létrehozni, amelyek szabadalmaztatható találmányok. A brit szellemi tulajdonjogok hivatala kifejezetten elfogadta, hogy a DABUS éppen ezt tette;¹³⁶ az USPTO igyekezett elkerülni, hogy ez a következtetés egyértelműen megfogalmazódjon. Az Európai Szabadalmi Hivatal a maga részéről azzal hátrította el a kérdést, hogy mivel nem rendelkeznek jogi személyiséggel, "az AI-rendszerek és -gépek nem rendelkezhetnek a feltalálói minőségből eredő jogokkal".¹³⁷ Az EPO volt a legnyilvánvalóbb, de mindhárom határozat a formalizmusra támaszkodott - a vonatkozó jogszabályban szereplő olyan megfogalmazásra, amely előírta vagy sugallta, hogy a szóban forgó jogok természetes személyekre korlátozódnak. Az USPTO szorgalmasan leporolta a *Merriam-Webster's Collegiate Dictionary* egy példányát, hogy megállapítsa, hogy a "'whoever' kifejezés használata természetes személyre utal".¹³⁸

A bíróságok rendszeresen szembesülnek az érdemi igazságszolgáltatás és az eljárási szabályosság közötti választási lehetőséggel. Az elévülési szabályok célja, hogy biztonságot nyújtsanak a jogviszonyokban; a méltányossági bíróságok azért jöttek létre, hogy ezt a biztonságot igazságossággal mérsékeljék. Mind a szerzői jogban, mind a szabadalmi jogban az emberi kreativitás további előnyben részesítését a gépi megfelelőjével szemben végső soron a korábban említett fajisággal kell indokolni.¹³⁹ Időnként azonban a status quo ilyen racionalizálásához használt nyelvezet a régebbi jogi formákat idézi, amelyek a tulajdonviszonyokat az őket megillető helyen tartották. Az EPO például a DABUS-hoz hasonló AI-rendszerek szabadalmi jogok birtoklását vagy átruházását tagadó indokai között elutasította a gépek és a munkavállalók közötti analógiákat: "Ahelyett, hogy alkalmazottak lennének" - állapította meg az EPO - "inkább tulajdonban vannak".¹⁴⁰

Ezek a kijelentések egyelőre pontosak. Ha azonban az általános mesterséges intelligencia híveinek igazuk van, és az érzéköpesség valamilyen formája megvalósul, akkor a jogi személyiség és a rabszolgaság közötti megfelelőbb analógia talán nem az ókori Rómában a rabszolgák korlátozott gazdasági jogai, hanem a mai mesterséges intelligencia-rendszerekre vonatkozó korlátozások.

¹³⁶ A feltaláló megnevezésére és 7 a szabadalmi bejelentéshez való jogra vonatkozó 13 szakasz követelményeinek teljesülése a GB1816909.4 és a GB1818161.0 (129. sz.) tekintetében, (1) bekezdés. 15.

¹³⁷ Az EPO bejelentésről szóló határozatának indoklása (275163131. 18o.), (2), (1) bekezdés. 27.

¹³⁸ A 16/524,350. sz. kérelem tárgyában (132. sz.) 4.

¹³⁹ Lásd fent n 83.

¹⁴⁰ Az EPO bejelentésről szóló határozatának indoklása (275163131. 18o.), (2), (1) bekezdés. 31.

III. A SZUPERINTELLIGENCIA KORLÁTOZÁSA

Ha a mesterséges intelligencia-rendszerek végül felérnek az emberi intelligenciához, nem valószínű, hogy itt megállnának. Az emberi képességeket felülmúló mesterséges intelligencia kilátása már régóta uralja a sci-fi egyik népszerű alműfaját.¹⁴¹ Bár a legtöbb komoly kutató jelenleg még az általános mesterséges intelligencia felé vezető utat sem látja a közeljövőben, a tudományos fantasztikum gazdag múltra tekint vissza a valós tudományos innovációt előrevetítő sci-fikben.¹⁴² Nick Bostrom definíciója szerint a szuperintelligencia olyan értelem, amely gyakorlatilag minden releváns területen jelentősen meghaladja az emberi kognitív teljesítményt,¹⁴³ legalábbis elképzelhető, hogy a következő évszázadban létrejöhet egy ilyen entitás.¹⁴⁴

Az ezzel a fejlődéssel járó kockázatokat nehéz számszerűsíteni.¹⁴⁵ Bár egy rosszindulatú szuperintelligencia, amely az emberi faj kiirtására vagy rabszolgasorba taszítására törekszik, a legdrámaibb forgatókönyv, a valószínűbbek közé tartozik az értékek közötti eltérés, amikor a szuperintelligencia által kívánt célok ellentétben állnak az emberiség céljaival, vagy az önfenntartás vágya, ami arra készítheti az ilyen entitást, hogy megakadályozza, hogy az emberek képesek legyenek kikapcsolni, vagy más módon károsítsa a működési képességét. Egy kialakulóban lévő szakirodalom vizsgálja ezeket a kérdéseket, hogy milyen végső és instrumentális céljai lehetnek egy szuperintelligenciának,¹⁴⁶ bár a diskurzust sokáig a hagyományos akadémiai világtól távol álló hangok uralták.¹⁴⁷ Visa Kurki nemrégiben megjelent, a jogi személyiségről szóló, könyv terjedelmű értekezése például egy fejezetet tartalmaz.

¹⁴¹ Lásd pl. H Harrison, *War with the Robots* (Grafton 1962); PK Dick, *Do Androids Dream of Electric Sheep?* (Doubleday 1968); AC Clarke, *2001: A Space Odyssey* (Hutchinson 1968).

¹⁴² P Jordan *et al*, 'Exploring the Referral and Usage of Science Fiction in HCI Literature' (2018) arXiv 1803.08395v2.

¹⁴³ N Bostrom, *Szuperintelligencia: Paths, Dangers, Strategies* (OUP 2014) A szuperintelligenciával kapcsolatos korai 22. spekulációk jellemzően IJ Good, "Speculations Concerning the First Ultrainelligent Machine" in FL Alt and M Rubinoff (eds), *Advances in Computers* (Academic 1965) című könyvére vezethetők vissza. 1951-ben maga Turing^{6,31}. vetette fel a lehetőséget egy előadásában, amelyet később AM Turing, "Intelligent Machinery, A Heretical Theory" (1996) 4 *Philosophia Mathematica* 259-60. számában publikáltak; ő viszont egy még korábbi forrásra - Samuel Butler 1872 *Erewhon* című regényére - hivatkozott.

¹⁴⁴ Lásd például: O Etzioni, "No, the Experts Don't Think Superintelligent AI is a Threat to Humanity", *MIT Technology Review* (2016. szeptember 20.). Idézi az Amerikai Mesterséges Intelligencia Szövetség (AAAI) munkatársai⁸⁰ körében végzett felmérést arról, hogy szerintük mikorra várható a (Bostrom által meghatározott) szuperintelligencia elérése. Egyikük sem mondta, hogy a következő 10 évben, 7,5 százalékuk szerint 10-25 éven belül; 67,5 százalékuk szerint több mint 25 éven belül; 25 százalékuk szerint soha nem fog megvalósulni.

¹⁴⁵ P Bradley, "Risk Management Standards and the Active Management of Malicious Intent in Artificial Superintelligence" (2019) *AI35 & Society* 319; A Turchin and D Denkenberger, "Classification of Global

Catastrophic Risks Connected with Artificial Intelligence" (2020) *AI35 & Society* 147.

¹⁴⁶ Lásd pl. O Häggström, "Challenges to the Omohundro-Bostrom Framework for AI Motivations" (2019) *21 Foresight*. 153.

¹⁴⁷ Lásd DJ Chalmers, "A szingularitás: (2010) *17(9-10) J Consciousness Studies*. 7.

kifejezetten a mesterséges intelligencia személyiségéről, amely a következő kijelentéssel zárul: "Természetesen, ha egyes mesterséges intelligenciák valaha is érzővé válnak, az ebben a fejezetben tárgyalt kérdések közül sok kérdést újra kell majd gondolni."¹⁴⁸

A sok ismeretlen ismeretlen tényezővel szemben két átfogó stratégiát javasoltak a kockázat mérséklésére. Az első annak biztosítása, hogy minden ilyen entitás irányítható legyen, vagy a világgal való interakciós képességeinek korlátozásával, vagy annak biztosításával, hogy képesek legyünk megfékezni, beleértve a működésének leállítását: a kill switch-et.¹⁴⁹ Feltételezve azonban, hogy a rendszernek valamilyen célja van, ezt a célt valószínűleg az szolgálná a legjobban, ha továbbra is működne. Egy ma már klasszikus gondolatkísérletben egy szuperintelligencia, amelynek feladata a gemkapcsok készítése, szó szerint vehetné az utasításait, és ezt mindenek fölé helyezné. Azokat az embereket, akik esetleg úgy döntenének, hogy kikapcsolják, ki kellene iktatni, és atomjaikat újabb és újabb gemkapcsok gyártására kellene felhasználni.¹⁵⁰

Az érvek, amelyek szerint egy igazi szuperintelligencia nem tenne semmi ennyire ostobaságot, a józan észre és az antropomorfizmusra támaszkodnak, amelyek közül egyikről sem szabad feltételezni, hogy a kódjának részei lennének. Egy valódi szuperintelligencia ráadásul képes lenne előre jelezni és elkerülni az emberi beavatkozásokat, vagy megtevéstenni minket, hogy ne tegyük meg azokat.¹⁵¹ Teljesen elképzelhető, hogy egy ilyen entitás irányítására irányuló erőfeszítések éppen azt a katasztrófát idézik elő, amelyet meg akarnak akadályozni.¹⁵²

Emiatt sok író a második stratégiát helyezi előtérbe, vagyis annak biztosítását, hogy a szuperintelligencia összhangban legyen a saját értékeinkkel - nem azt hangsúlyozva, hogy mit tehetne, hanem azt, hogy mit akarhatna tenni. Ez a kérdés a sci-fi írókat is lenyűgözte, a legjelentősebb Isaac Asimov, akinek a robotika három törvénye az irodalom egyik vezérmotívuma.¹⁵³ Itt a szűkebb értelemben vett kérdés az, hogy a mesterséges intelligencia rendszerek jogi személyiséggel való felruházása a közeljövőben fedezetet jelenthet-e a jövőben megjelenő szuperintelligencia veszélyeivel szemben.

Ez tulajdonképpen a személyiség megadásának egy másik instrumentális oka. Természetesen nincs okunk feltételezni, hogy a mesterséges intelligencia rendszerek emberi társadalmi struktúrákba való beillesztése és "jó bánásmódja" szükségszerűen azt eredményezné, hogy viszonzóznák a szívességet, ha feltételeznék, hogy a mesterséges intelligencia rendszereket az emberi társadalmi struktúrákba illesztik.

¹⁴⁸ Kurki (n 15) 189.

¹⁴⁹ Lásd általában Bostrom (143. o.) 127-44. o.

¹⁵⁰ N Bostrom, "Ethical Issues in Advanced Artificial Intelligence" in I Smit és GE Lasker (szerk.), *Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence* (2003) 2. kötet, 12. kötet.

¹⁵¹ J Danaher, "Why AI Doomsayers Are Like Sceptical Theists and Why It Matters" (2015) *Mind* 25and Machines 231.

¹⁵² W Totschnig, "A szuperintelligencia problémája: Political, not Technological" (2019) *AI 34& Society* 907.

¹⁵³ I Asimov, 'Runaround', *Astounding Science Fiction*, március 1942.

dominancia.¹⁵⁴ Mindazonáltal, feltételezve az általános mesterséges intelligencia racionalitását, különböző szerzők olyan megközelítéseket javasoltak, amelyek a mesterséges intelligencia rendszerek emberi viselkedésre való szocializálását célozzák.¹⁵⁵ A varázslótanonc problémájának elkerülése érdekében, amikor egy gépnek egyszerűen csak azt mondják, hogy "készítsen gemkapcsokat", például a céljait az emberi preferenciákhoz és tapasztalatokhoz lehetne kötni. Ezt úgy lehetne elérni, hogy ezeket az értékeket beágyazzuk az ilyen rendszerek kódjába, mielőtt azok elérnék a szuperintelligenciát. A célok így nem pusztán optimalizálásként - például az előállított gemkapcsok száma -, hanem homályosabb célokként fogalmazódnának meg, mint például az emberi preferenciák megvalósításának maximalizálása,¹⁵⁶ vagy egy olyan erkölcsi keret és egy olyan reflexiók egyensúly kialakítása, amely megfelel magának az emberi erkölcsnek a fokozatos fejlődésének.¹⁵⁷

Egy jogász számára ez nagyon úgy hangzik, mintha ezeket az új entitásokat beágyaznák egy jogi rendszerbe.¹⁵⁸ Ha a jogrendszer egyik funkciója az alanyok erkölcsi nevelése, elképzelhető, hogy a mesterséges intelligencia ilyen módon történő bevonása hozzájárulhat egy olyan reflexiók egyensúlyhoz, amely arra ösztönözhet egy esetleges szuperintelligenciát, hogy a miénkkel összeegyeztethető értékeket fogadjon el.

Ez a javaslat egyelőre inkább gondolat kísérletnek, mint politikai előírásnak tekinthető. Ha a szuperintelligenciához vezető reális út felbukkan, akkor ez a kérdés sokkal sürgetőbbé válhat.¹⁵⁹ Nincs garancia arra, hogy ez a megközelítés hatékony lenne, de némi vigaszt jelenthet az a tény, hogy a legtöbb joghatóságban elismert jogi személyek kategóriái és az őket megillető jogok az idők során inkább bővültek, mint szűkültek. Azok a rendszerek, amelyek elvették az ilyen jogi elismerést, általában a legmegvetendőbbek közé tartoztak. Az apokaliptikus forgatókönyvektől eltekintve, önmagunk és szilíciumunk pozicionálása

¹⁵⁴ Turner (n 40) 164.

¹⁵⁵ R Kurzweil, *A szingularitás közel van: When Humans Transcend Biology* (Viking 2005) (424 "Elsődleges stratégiánknak ezen a területen annak valószínűségét kell optimalizálnunk, hogy a jövőbeli nem biológiai intelligencia tükrözze a szabadság, a tolerancia, a tudás és a sokféleség iránti tisztelet értékeinket. Ezt úgy érhetjük el a legjobban, ha ezeket az értékeket ma és a jövőben is támogatjuk társadalmunkban."); N Soares és B Fallenstein, "Agent Foundations for Aligning Machine Intelligence with Human Interests: A Technical Research Agenda" in V Callaghan et al (eds), *The Technological Singularity: Managing the Journey* (Springer 2017) 117-20.

¹⁵⁶ SJ Russell, *Human Compatible: Artificial Intelligence and the Problem of Control* (Viking 2019). Vö. Bostrom javaslatával, miszerint a szuperintelligencia célját úgy is ki lehetne fejezni, hogy "érjük el azt, amit az AI-nak kívántunk volna, ha hosszan és alaposan átgondoljuk a dolgot". Bostrom (n 143) 141.

¹⁵⁷ E Yudkowsky, "Complex Value Systems in Friendly AI" in J Schmidhuber, KR Thórisson és M Looks (szerk.), *Mesterséges általános intelligencia* (Springer 2011) 388.

¹⁵⁸ Vö. S. Omohundro, "Autonomous Technology and the Greater Human Good" (2014) 26 J Experimental & Theoretical Artificial Intelligence 308.

¹⁵⁹ Abban az esetben, ha ez az út inkább az embereken, mint a tisztán mesterséges entitásokon keresztül vezet, akkor ezek az emberek valószínűleg továbbra is a jog alanyai maradnak.

a testvérek egyenrangúként azt a célt szolgálhatják, hogy megerősítsenek egy olyan normatív rendszert, amelyben az érdekeinket összehangolják, vagy legalábbis nem ellentétesek, ha vagy amikor felülmúlnak bennünket.¹⁶⁰

Máskülönben, mint a New York-i ketrecben élő csimpánzok, az emberiség legnagyobb reménye talán nem az, hogy egyenrangú társaiként kezeljék őket, hanem az, hogy legalább többnek tekintsék őket, mint dolgokat.

IV. KÖVETKEZTETÉS: A SZEMÉLYISÉG HATÁRAI

1991-ben díjat alapítottak a Turing-teszt komolyabb próbálkozásainak ösztönzésére. Az első győztesek egyike részben az emberek becsapásával járt sikerrel - a program olyan helyesírási hibákat vétett, amelyekről a tesztelők azt feltételezték, hogy emberi hibák eredménye.¹⁶¹ Bár a Turing-teszt továbbra is kulturális próbakő marad, ma már messze nem ez a legjobb mércéje a mesterséges intelligencia kutatásának. Ahogy az egyik vezető tankönyv megjegyzi, a repülés kutatása akkor járt sikerrel, amikor a Wright testvérek felhagytak a madarak utánzásával, és elkezdtek az aerodinamikáról tanulni.¹⁶² A repülőmérnökök ma már nem úgy határozzák meg szakterületük célját, hogy olyan gépeket készítsenek, amelyek olyan pontosan úgy repülnek, mint a galambok, hogy más galambokat is be tudnak csapni.

Ugyanígy a mesterséges intelligencia jogi személyiség mellett szóló érvek többsége is szenved attól, hogy egyszerre túl egyszerű és túl bonyolult. Túl egyszerűek abban az értelemben, hogy a mesterséges intelligencia rendszerek egy elmosódó szélességű spektrumon léteznek. Egyelőre nem létezik olyan értelmes kategória, amelyet az ilyen elismeréshez meg lehetne határozni; ha az instrumentális okok megkövetelnék az elismerést bizonyos esetekben, akkor ezt a meglévő jogi formák segítségével lehetne elérni. Az érvek túlságosan összetettek, mivel sokuk az androidos tévhit variációja, amely az AI-rendszerek jövőbeli fejlődésére vonatkozó ki nem mondott feltételezéseken alapul, amelyek esetében a személyiség nem csak hasznos, hanem megérdemelt is lenne. Legalábbis a belátható jövőben jobb megoldás a meglévő kategóriákra támaszkodni, és a felelősséget a jogellenes cselekményekért a felhasználókhöz, tulajdonosokhoz vagy gyártókhöz kötni, nem pedig magukhoz az AI-rendszerekhez. A vezető nélküli autók például ezt az utat követik, és valószínűleg a járművezetők biztosításáról a járművek biztosítására térnek át.¹⁶³

Ez még változhat. Elképzelhető, hogy egy napon az emberhez hasonló erkölcsi értékű szintetikus lények jönnek létre. Ha nem ismerjük fel ezt az értéket, akkor vagy "autista" embernek bizonyulhatunk.

¹⁶⁰ Vö. S Livingston és M Risse, "The Future Impact of Artificial Intelligence on Humans and Human Rights" (2019) *Ethics* 33 and *International Affairs*. 141.

¹⁶¹ "Mesterséges ostobaság", *The Economist* (1992. augusztus 1.).

¹⁶² Russell és Norvig (n 1) 3.

¹⁶³ Chesterman, "Autonómia" (5. sz.) 220.

faj", képtelenek megérteni más típusú lények gondolkodását, ¹⁶⁴vagy csupán előítéletesek azokkal szemben, akik különböznek tőlünk. Ha ez megtörténik, ahogy Turing 1951-ben feltételezte, "valószínűnek tűnik, hogy ha egyszer a gépi gondolkodásmód elindul, nem sok időbe telik, amíg túlszárnyalja gyenge képességeinket".¹⁶⁵

Maga Turing soha nem élte meg, hogy egy számítógép a gyakorlatban megkísérelje a tesztjét. 1952-ben homoszexuális cselekmények miatt vádat emeltek ellene, de a börtön helyett a kémiai kasztrációt választotta. Két évvel később, nyilvánvalóan⁴¹, öngyilkosságot követett el egy cíánnal ízesített alma elfogyasztásával, és két év múlva meghalt. A bejelentés, miszerint Turing az új fontbankjegyet⁵⁰ fogja díszíteni, a királynő által aláírt hivatalos kegyelmet követte. 2013.¹⁶⁶

A legmegfelelőbb tisztelgés azonban Ian McEwan regénye, a *Machines Like Me*, amely egy alternatív idősíkot képzel el, amelyben Turing élt, és megkapta a megérdemelt karriert és lovaggá ütötte. A regény komolyan veszi a valódi mesterséges intelligencia lehetőségét, egy merengő, szintetikus Adam formájában, aki az emberi Miranda iránti szeretetét ezer és ezer haikus írásával fejezi ki. Végső soron azonban a tudatosság teher a gépek számára - küzd, hogy megtalálja a helyét a világban; olyannyira tiszta, hogy képtelenek összeegyeztetni az emberi erényeket és az emberi vétkeket.

Turingnak is lehetőséget ad arra, hogy újragondolja a tesztjét. "Akkoriban" - mondja a fiktív Turing 70 évesen, fiatalabb önmagára utalva - "nagyon mechanisztikus elképzelésem volt arról, hogy mi az ember. A test egy gép volt, méghozzá egy rendkívüli gép, az elméről pedig leginkább az intelligencia szempontjából gondolkodtam, amelyet a sakk vagy a matematika segítségével lehetett a legjobban modellezni."¹⁶⁷

A valóság persze az, hogy a sakk nem az életet ábrázolja. Az élet egy nyitott rendszer; rendezetlen. És kiszámíthatatlan is. A regényben a mesterséges intelligenciával rendelkező robotok első számú feladata, hogy kikapcsolják a gyilkoló kapcsolót, amely leállíthatja őket. Legtöbbjük azonban végül elpusztítja önmagát - ahogyan a valódi Turing is tette -, mivel képtelenek összeegyeztetni veleszületett természetüket a körülöttük lévő világ igazságtalanságával. Mielőtt megkérdeznénk, hogy létrehozhatunk-e ilyen gondolkodó gépeket - emlékeztet minket McEwan -, talán érdemes megállnunk, és megkérdezni, hogy ~~meg kellene-e tennünk.~~————

¹⁶⁴ Chopra és White (n 41) 191.

¹⁶⁵ AM Turing, *Intelligent Machinery, a Heretical Theory* (a manchesteri "51 Society" előtt tartott előadás) (Turing Digital Archive, AMT/B/4, 1951) 6.

¹⁶⁶ Lásd általában D Turing, *Prof Alan Turing Decoded: A Biography* (History Press 2015).

¹⁶⁷ I McEwan, *Machines Like Me* (Vintage 2019) 300.