

Eszteri Dániel

A mesterséges intelligencia fejlesztésének és üzemeltetésének egyes felelősségi kérdései

Vajon ha egy mesterséges intelligencia, amely első sorban azért íródott, hogy bizonyos mértékben helyettesítse az emberi gondolkodást és önállóan hozzon döntéseket, működése során hibát követ el, esetleg kárt okoz, ki tehető azért felelőssé?

A csúcstechnológia fejlődése által generált problémák számos újszerű jogi kérdést is felvetnek, hiszen annak irányait rendkívül nehéz megjósolni, és hogy az vajon milyen utakat fog a jövőben bejárni. Az 1970-es, 80-as évek science-fiction filmjeiben és novelláiban bemutatott eszközök mára a mindennapi élet részeivé váltak nem egy esetben. Az internet által megnyitott virtuális térben az emberi felhasználók sok esetben mesterséges intelligenciák által irányított szoftverekkel kommunikálnak.

A tanulmány első sorban azért íródott, hogy a mesterséges intelligencia funkcionalitással rendelkező szoftvereket érintő jogi problémákat elemezze és választ adjon arra a kérdésre, hogy vajon az azok működéséből eredő hibákért, illetve azok cselekedeteiért ki tehető felelőssé. A tanulmányban kitérek a mesterséges intelligenciák általános jellemzőire és a társadalmunkra gyakorolt hatásaira, a mesterséges intelligencia és a jog kapcsolatának vázolására, érintve a vonatkozó törvényi szabályokat és azok kritikáját. A fő megválaszolásra váró kérdés az ilyen létezők által okozott károkért való felelősség vizsgálata lesz, mind magánjogi, mint büntetőjogi oldalról.

1. A mesterséges intelligenciák általános jellemzői és felosztásuk

Az informatika fejlődése nem egy olyan jelenséget hívott életre az elmúlt évtizedekben, ami az emberi társadalom berendezkedését alapvetően alakította át. Ilyenre lehet példa a könyvek digitalizálása, a kommunikáció felgyorsulása és általában az információs társadalom fejlődése. Az egyik legmeghökkenőbb és épp emiatt sok futurologust, írókat és művészt megihlető jelenség azonban kétség kívül a mesterséges intelligenciák működése. A mesterséges intelligenciát is tartalmazó szoftverek már nem csupán eszközök az ember kezében, mivel azok képesek akár önálló döntéseket is hozni azon feladatok ellátása céljából, amik elvégzésére írták őket.

A mesterséges intelligencia használata manapság már kilépett a tudományos kísérletezés kereteiből és a mindennapok részévé vált. Ilyen programok mutatják meg az óceánjáró hajók legénységének, hogy a tengerfenéken melyik objektum természetes vagy mesterséges képződmény, megjósolják, hogy a piaci mechanizmusok milyen mértékben hatnak az értékpapírtőzsde működésére,¹ megmutatják, hogy milyen úton tudunk a legrövidebb idő alatt és a legkevesebb benzinköltséggel eljutni a szomszéd városba, de ezek irányítják a számítógépes játékprogramok ellenfeleit is, amelyeken az emberi játékos küldetése felülkerekedni.

Mivel a legtöbb mai jogrendszerben nincs (köztük a magyar jogban sem) meghatározva a mesterséges intelligencia fogalma, érdemes kitérni először arra, hogy a filozófia és a technika fejlődésének szempontjából történetileg mit is tekintettek annak az egyes szerzők.

Az Oxford számítástechnikai értelmező szótár szerint a mesterséges intelligencia a számítástudománynak az a területe, amely emberi intelligenciát igénylő feladatokat megoldó számítógépes programok készítésével foglalkozik.²

STUART RUSSEL és PETER NORVIG a mesterséges intelligencia fogalmának filozófiai fejlődése szempontjából négy féle irányzatot különböztetett meg, amelyek a következők.³

1. *Emberi módon gondolkodó rendszer:* Az emberi elme működését és a megismerést modellező rendszereket tekinti ez az irányzat mesterséges intelligenciának.
2. *Emberi módon cselekvő rendszer:* Ez a megközelítés ALAN MATHESON TURING matematikus nevéhez kötődik, aki a róla elnevezett Turing-teszt alapján az emberi viselkedést állította az intelligencia fő kritériumának, és az elérendő célnak.

3. *Racionálisan gondolkodó rendszer*: Ez az irányzat az emberi gondolkodásnál valamilyen értelemben tökéletesebb, racionálisabb gépek és szoftverek megalkotásában látja a mesterséges intelligencia fejlesztés célját.
4. *Racionálisan cselekvő rendszer*: A modern informatikai tudományok megközelítése, amely nem tűzi ki azt célul, hogy a kialakított rendszerek a szó klasszikus értelmében gondolkodjanak, vagy imitálják az emberi viselkedést, csak azt, hogy minél racionálisabban viselkedjenek (pl. egyértelműen tudják diagnosztizálni a betegségeket, jelezzék előre a természeti katasztrófákat stb.). A fejezet későbbi pontjaiban, az MI gyakorlati alkalmazásával kapcsolatban ezt a felfogást tekintem irányadónak.

JOHN R. SEARLE definíciója szerint továbbá különbséget tehetünk gyenge és erős mesterséges intelligencia között. *Gyenge MI*-nek Searle azon rendszereket tekinti, amelyek úgy cselekszenek, mintha intelligensek lennének, de ennek ellenére arról már nincs információ, hogy a rendszer valóban rendelkezik-e elmével, vagy sem. Az *erős MI* ezzel szemben az olyan rendszerekre értendő, amelyek valóban gondolkodnak, önálló tudatuk van.⁴

A filozófia elsősorban arra kereste a választ, hogy vajon létrehozható-e az *erős MI*, tehát hogy szimulálható-e az emberi elme kognitív működése, rendelkezhet-e egy mesterséges entitás önálló tudattal. Több érv is született a történelem során, mind pro, mind kontra. Ezek az elméletek az emberi elme működésével kapcsolatban a gondolkodás és nyelv viszonyának szánnak kulcsszerepet a kérdés megválaszolásában.⁵

Már PLATÓNnál megjelent a probléma a gondolkodás és a nyelv kapcsolatáról. Számára úgy tűnik, hogy a gondolkodás és a nyelv azonos egymással, mivel a gondolat és a beszéd ugyanannak az aktusnak a belső és külső oldala.⁶

Ez a feltevés más gondolkodóknál is megjelenik (pl. HUMBOLDT), mely szerint a nyelv és a gondolkodás azonos, de legalábbis szoros kapcsolatban van egymással.

ARISZTOTELÉSZ híres szillogisztikus módszere szerint – amely a helyes gondolkodás szabályait foglalja magában –, a nyelv tanulmányozása nem a szavak tanulmányozását jelenti, hanem a gondolkodását, amelynek a szavak a jelei.

DESCARTES *Értekezés a módszerről* című művében veti fel azt a problémát, hogy ha léteznének külsőleg és viselkedésben is teljesen az emberre emlékeztető gépek, akkor hogyan tudnánk megállapítani, hogy mégsem egy organikus emberi lényel, hanem annak szintetikus másával van dolgunk. Első érve, hogy szerinte a gépeknek sosem lehet önálló tudatuk, mivel a nyelvet nem a belső gondolataik kifejezésére használják, hanem egyes külső hatásokra (pl. emberekkel való interakció) történő reagálásra. Második érve, hogy noha elképzelhető, hogy a gépek egyes feladatokat ugyanolyan jól, vagy még jobban el tudnak végezni, mint az emberek, azonban másokban okvetlenül csődöt mondanak, mivel nem tudatosan cselekszenek, hanem szerkezetük révén.⁷

2. A mesterséges intelligencia posztmodern elméleti háttére

2.1 A Turing-teszt

Mindezek a történelmi előzmények jelzik, hogy az embert már évezredek óta foglalkoztatja a kérdés, hogy vajon lehet-e alkotni mesterséges úton olyan lényeket, amelyek képesek az emberi, vagy más organikus lények viselkedéshez hasonlóan önállóan döntéseket hozni és cselekedni. Mindezen eszmék kétségkívül megalapozták a mesterséges intelligencia kutatás egyik legjelentősebb művét, ALAN MATHESON TURING *Computing Machinery and Intelligence* című cikkét, amelyet 1950-ben jelentetett meg. Descartes érvelésével némileg ellentétesen Turing arra az álláspontra helyezkedik, hogy abban az esetben, ha képesek lennének egy olyan megfelelően programozott számítógépet létrehozni, amely tetszőleges témáról úgy tudna beszélgetni egy emberi féllal, hogy az nem tudná megállapítani, hogy egy számítógéppel vagy egy másik élő emberrel folytat eszmecsere, akkor feltételeznünk kéne, hogy a gép intelligens. A Turing-teszt néven elhíresült kísérletben egy kérdezőbiztos feladata lenne a fenti következtetés megállapítása, úgy, hogy egy billentyűzet és monitor segítségével tartaná a kapcsolatot a másik szobában lévő tesztalánnyal, akinek kérdéseket kellene feltennie, majd az adott válaszok alapján kéne kitalálnia, hogy emberi lényel, vagy egy számítógéppel beszélgetett-e.⁸

Turing-tesztje tehát kikerüli azt, hogy az intelligenciát és a gondolkodást feltétlenül összemoszuk egymással. Egyszerűen azt feltételezi, hogy a teszt elég nehéz ahhoz, hogy aki (ami) át tud menni rajta, azt intelligensnek kell tekinteni.

A kognitív tudományok azzal a feltevessel számolnak, hogy az emberi intelligencia különböző számításokon alapul, így az emberi elmét elvileg lehetséges modellezni egy számítógépes programon keresztül.⁹ A probléma először HOBBS-nál jelent meg a következő kijelentésében: „Racionalitás alatt számításokat értek”.¹⁰

A XX. század második felében a filozófiában ezzel szemben inkább azok az álláspontok kerültek előtérbe, amely szerint az emberi intelligenciát nem lehet megfeleltetni a Turing-teszt által vázolt intelligencia kritériummal, mivel mindig lehetséges olyan állításokat generálni, amelyet egy formális rendszer nem tud bizonyítani. Ezt az álláspontot az irodalomban többek között KURT GÖDEL, JOHN LUCAS és ROGER PENROSE képviseli.¹¹ Elméleteik bemutatására ehelyütt azonban, azok szerteágazó volta miatt bővebben nem térek ki.

2.2 A kínai szoba argumentum

JOHN R. SEARLE a Turing-teszt alapján felvetetteket, valamint az intelligencia és a számítási képesség viszonyát az úgy nevezett „kínai szoba” elmélettel kérdőjelezte meg. Az elmélet tesztalánya egy emberi lény, aki egy szobába van bezárva. Kívülről egy ablakon keresztül különböző kérdéseket kap papírlapon, amik kínai nyelven íródtak. A szobába zárt tesztalany nem tud kínaiul, azonban van nála egy szótár, amiből ki tudja keresni a különböző leírt szavak jelentését, csupán a kínai írásjelek formája alapján. Ezen felül a könyv különböző szabályokat is tartalmaz az írásjelek megformálásával kapcsolatban, tehát egyszerre szótár és egyfajta nyelvtankönyv. A kínaiul íródott kérdéseket a szobában lévő fél a szótár segítségével lefordítja, majd megfogalmazza rá a válaszokat is szintén ezen a nyelven. A szobán kívüli emberek ebben az esetben tehát egyfajta Turing-tesztben vesznek részt a tesztalanyal együtt, hiszen úgy hiszik, hogy a bent lévő fél ismeri a kínai nyelvet. A szobában lévő tesztalany azonban igazából nem tud kínaiul, csak különböző a kínai írásjelek formájával kapcsolatos instrukciókat követ. Searle tehát azt állítja, hogy ebben az esetben sem a tesztalany, sem a könyv nem érti a használt nyelvet, mégis képesek rajta kommunikálni másokkal. Így kell szerinte elképzelni azt a szituációt is, amikor a számítógép kommunikál egy emberrel. Abban az esetben a tesztalany a számítógép (hardver), a könyv pedig a program (szoftver), amelyek együttesen próbálnak intelligens kommunikációt imitálni, holott önmagukban nem intelligensek.¹²

2.3 Az agyprotézis kísérlet

SEARLE és TURING feltevései szerint tehát a mesterséges intelligencia nem más, mint szimuláció, amittől nem feltétlenül kell elvárni, hogy valóban gondolkodjon. Ezekkel az álláspontokkal szemben hozza fel HANS MORAVEC az úgy nevezett agyprotézis kísérletet, melynek érvelése azon a feltevésen alapul, hogy neurobiológia már odáig fejlődött, hogy tökéletesen feltérképezte az agy összes neuronját, valamint az azok közötti összeköttetéseket is egytől-egyig ismeri, és ez alapján tudjuk, hogy adott neuron adott bemenetre milyen kimenetet produkál. Tegyük fel, hogy a tudomány képes e viselkedéseket utánozó „mű neuronokat” építeni. Az argumentum szerint egy műtét során folyamatosan kicserélik egy ember agyának valódi neuronjait szintetikus neuronokra. Ekkor MORAVEC szerint az agy funkcionálisan ugyanúgy működik, mint előtte, azonban a kérdés, hogy vajon mi történt a műtétnek alávetett ember tudatával: megmaradt, vagy eltűnt a műtét közben? MORAVEC szerint a kísérleti alany tudata tökéletesen megmaradna, mivel a neuronok funkcionális tulajdonságai az egyedüli lényeges tényezők annak szempontjából.¹³ SEARLE ezzel nem ért egyet, szerinte a kísérleti alany tudata fokozatosan eltűnne a kísérlet alatt.¹⁴

Az agyprotézis kísérlet mögötti filozófia elevenedik meg a Ghost in the Shell című japán animációs filmben is, amelynek főhősnője, Kuszanagi őrnagy ugyan emberként született, azonban élete során a folyamatos mechanikus implantátumok beültetései miatt agyszövetein kívül szinte teljes egészében robottá vált, habár kinézete megmaradt emberinek. A hősnő filmbéli monológja alkalmával kifejti, hogy nem tudja egyértelműen eldönteni a kérdést, mely szerint még saját emberi tudata, vagy már esetleg egy azt tökéletesen szimuláló szoftvert működik az agyában.¹⁵

2.4 Mesterséges intelligencia szimulációk a gyakorlatban

A fenti elméletek alapján a Turing-teszt tehát csak arra alkalmas minden kétséget kizáróan, hogy olyan rendszereket lehessen azonosítani vele, amelyek az emberi viselkedést utánozzák és gyenge MI-nek tekinthetők. Ha egy számítógépes program esetleg át is megy a Turing-teszten – tehát emberi módon válaszol a neki feltett kérdésekre –, az még nem garancia arra, hogy rendelkezik önálló elmével és valóban gondolkodik, ami az erős MI legfőbb jellemzője. Az erős MI a technikai jelen állása szerint még nem létezik, és ki tudja, hogy egyáltalán belátható időn belül sor fog-e kerülni létrehozására. A gyenge MI-re azonban számtalan példát találunk a mindennapi életben, ennek tulajdonságait legjobban talán a gyakorlatban is működő néhány intelligens viselkedést szimuláló szoftver tudja átadni. A gyenge MI-eket azonban a legtöbb esetben nem is az emberi viselkedés imitálására, hanem feladatok racionális végrehajtására tervezik, ami a modern informatikai tudományok általános elfogadott megközelítése a mesterséges intelligenciák fejlesztése felé.

A szoftverek legegyszerűbben leírva úgy működnek, hogy fogadják a külső világ felől érkező bejövő, vagy szándékosan betáplált jeleket (inputok), majd ezeket előre megírt algoritmusok alapján feldolgozzák és végül kimenő jelekként (outputok) jelzik vissza az eredményt. Amíg az emberi elme fel tudja fogni, hogy adott input milyen algoritmuson megy keresztül és az milyen outputot fog szülni adott esetben, addig a számítógépes szoftverek a jogi felelősség szempontjából nem érdemelnek több figyelmet, mint egy egyszerű zsebszámológép. Azonban ahogy a sebesség, komplexitás és számítóképesség eléri azt a szintet, hogy az emberi megfigyelő már nem képes előre megjósolni az inputra adott választ, a történet érdekes fordulatot vehet nem csak technikai, de jogi szempontból is: kiszámíthatatlanná válik a gépek viselkedése az emberi megfigyelő számára: elkezdnek emberi, vagy akár emberfeletti képességeket tanúsítani és gyenge MI-vé válnak.¹⁶

Az egyik legjobb példa a gyenge MI-re és azok működési elveire a sakk játékprogram. Tudjuk, hogy léteznek olyan sakkjáték szimulátor szoftverek, amelyek képesek akár a legjobb sakk mestereket is legyőzni. Ilyen például az IBM cég Deep Blue nevű szoftvere, amely arról vált híressé, hogy 1997-ben sikerült megvernie a világ egyik legjobb sakkozójának tartott GARRY KASPAROVOT.¹⁷ Ennek ellenére mégsem lehet azt mondani, hogy Deep Blue ugyanolyan módon viszonyul és érti a játékot, mint Kasparov, annak ellenére, hogy kétségkívül mindketten lenyűgöző játéktudásról tesznek tanúbizonyságot. Deep Blue lépéseit szinte képtelenség előre megjósolni, ahhoz túl gyors, túl hatékony, túl komplex, az is előfordul, hogy – az emberre kísértetiesen hasonló módon – hibázik is, sőt tanul ellenfele stílusából.¹⁸ Azonban amikor a szoftver felülkerekedik az emberi játékoson, akkor nem a szoftvert programozó ember győz, hanem „maga Deep Blue”. A szoftverfejlesztőnek valószínűleg esélye sem lenne Kasparov ellen. Deep Blue tehát intelligensen játszik, attól függetlenül, hogy felfogja-e mit csinál egyáltalán, hiszen ez a körülmény a sakkjáték szempontjából teljesen irreleváns.¹⁹

A sakkprogramot azonban csak és kizárólag a sakk játszása szempontjából tekinthetjük valamilyen szinten intelligens entitásnak, mivel az semmi máshoz nem ért. A sakk szabályain és azok alkalmazásán kívül számára nem létezik semmi a világon, a játékban felmerülő problémákon kívül más természetű kihívásokat képtelen lenne értelmezni. Ez azonban önmagában nem probléma, hiszen a sakkprogram egy tipikusan racionálisan cselekvő rendszer és ennél nem is várunk el tőle többet. A mesterséges intelligencia legtöbb kutatója is azt a megközelítést részesíti előnyben, ha a rendszerek gyakorlati problémamegoldó képességére összpontosítunk, nem pedig az emberek utánzására való képességére.

3. A mesterséges intelligencia és a jog kapcsolata

3.1 A mesterséges intelligencia és a robotika társadalmi hatásának alapjai

A „teremtett értelem” társadalomra gyakorolt hatásának problémája ősidők óta foglalkoztatja az embereket. Az egyes történetek mögött általában az a mélyen gyökerező ösfélelem lapul meg, hogy az önálló tudatra ébredő mesterséges lény akár el is pusztíthatja a megteremtőjét. Ezen pesszimista irányzatok gyökerei a XX. század előtti szépirodalomban és folklórban is megtalálhatóak.

Jó példa erre MARY SHELLEY Frankensteinje, amelyben az emberi testrészekből összeállított, tudatára ébredő, „eredendő bűn nélküli” lény magányosságára hivatkozva ráveszi teremtőjét, a doktort, hogy egy női társat alkosson számára. Amikor azonban a főhős később megszegi a lénynek tett ezen ígéretét és félbeszakítja munkáját, a lény bosszút esküszik és fellázad teremtője ellen, ennek folyamányaképpen pedig először egyik legjobb barátját, majd menyasszonyát is meggyilkolja.²⁰

BARTHELMESS és FURBACH a koblenzi egyetem kutatói szerint a fent említett mondák vallási alapja mindkét esetben az, hogy az embereknek tiltott, hogy istenként, így jelen esetben teremtőként viselkedjenek.²¹ A teremtője ellen fellázadó mesterséges lény archetípusa ezért tulajdonképpen Isten haragját szimbolizálja, ami figyelmeztetni hivatott arra az embert, hogy a teremtő akaratával ellentétes cselekvés büntetést von maga után. Ez az archetipikus történet a XX. és XXI. században is számos olyan sci-fi elbeszélés alapját adta, mint például a 2001: Űrodüsszeia, a Mátrix, vagy a Szárnyas Fejvadász, hogy csak a legismertebbeket említsük.

Természetesen az emberek és a mesterséges lények közötti konfliktus nem csak az irodalmi fikció szintjén jelent meg. Az ipari forradalom alatt a gépektől való rettegés szülte például a Ludd tábornokként elhíresült NED LUDD vezette géprombolók mozgalma az 1810-es években. Az álarcos gépromboló munkások attól való félelmükben, hogy a szövőgépek kiszorítják az emberi munkaerőt az üzemekből, megrohanták a pamut- és gyapjúfonodákat és tönkretették a gépeket. Ennek hatására az angol parlament 1812-ben elfogadott egy olyan törvényt, amely a géprombolást halálbüntetéssel rendelte büntetni. A büntetési tételt egy évvel később 7-14 évig tartó deportációra enyhítették.²² A technikai fejlődés tagadása és a géprombolás gondolatának továbbvitele jelenik meg továbbá olyan XX. századi anarcho-primitivista írók munkáiban, mint a levélbombs támadásairól elhíresült THEODORE J. KACZYNSKI.²³

A mesterséges intelligencia fejlesztés pesszimista jövőképe a mai napig élénken él nem csak a laikusok, hanem a tudományos élet képviselői között is. A legutóbb olyan nevek is figyelmeztetéseket fogalmaztak meg, arra nézve, hogy az MI helytelen irányokba történő fejlődése akár az emberi faj végét is jelentheti, mint a Microsoft alapítva BILL GATES, vagy STEPHEN HAWKING a Cambridge Egyetem elméleti fizikus professzora.²⁴

A pesszimista elméletek alapját elsősorban az úgy nevezett „technológia szingularitás” problémája adja. RAY KURZWEIL szerint a szingularitás egy jövőbeli korszak, melyben a technológiai változás üteme olyan gyors lesz, a hatása pedig olyan mély, hogy az emberi élet visszafordíthatatlanul átalakul.²⁵ A szingularitás hatására megjelenő emberfeletti intelligencia miatt a technológiai fejlődés és ezzel összefüggésben a társadalmi változások oly mértékben felgyorsulnak, hogy a környezet megváltozását a szingularitás előtt élők képtelenek felfogni. A technológia szingularitás, avagy intelligenciारobbanás utáni eseményeket az elmélet szerint tehát a jelenlegi jövőképeinkkel képtelenek vagyunk megjósolni. A szingularitás nyomán megjelenő szuperintelligens mesterséges lények pedig könnyen kiszoríthatják az embert a létezésből. Az esemény bekövetkezését a jövőkutatók többnyire a XXI. század harmadik évtizedére jósolják.²⁶

Más, optimistább elméletek (pl. I.J. GOOD, MORAVEC) szerint az embereket leigázó MI víziója az ismeretlentől való rettegésből fakad, csakúgy, mint korábban a szellemektől, vagy boszorkányoktól való félelem. Ha a mesterséges intelligenciákat megfelelően, azaz olyan ágensekként tervezik, amelyek a gazdáik céljait teljesítik, akkor a jelenlegi tervezés lépésenkénti előrehaladásából származó MI-k szolgálni fognak, nem pedig leigázni. Az emberek azért használják agresszíven az intelligenciájukat, mert a természetes kiválasztódás miatt velük született agresszivitással rendelkeznek. De a gépek, amiket magunk építünk, nem születnek agresszívnak, hacsak nem döntünk úgy, hogy ilyenek építjük őket. Másrészt viszont lehetséges, hogy a számítógépek meghódítanak minket abban az értelemben, hogy szolgálatukkal elengedhetetlenné válnak, csakúgy, mint az autók meghódították ebben az értelemben az iparosodott világot.²⁷

A fentiek alapján is világosan látszik, hogy a mesterséges értelem ideális társadalmi és ezzel összefüggésben jogi szabályozása rendkívül fontos kérdés, és annak jelentősége minden bizonnyal egyre fontosabb lesz az idő előrehaladásával. A jogi szabályozás kapcsán felvetődő kérdéseket az alábbiakban tekintem át.

3.2 A mesterséges intelligencia és a robotika fogalmainak elhatárolása

A mesterséges intelligencia fejlesztése mérnöki és informatikai szempontból élesen elhatárolja egymástól az intelligens viselkedést imitáló szoftver (ágens) és a fizikai beavatkozó szervekkel is ellátott robot fogalmát. A két fogalmat azért szükséges elhatárolni, mivel számos olyan mesterséges intelligencia szimuláció létezik, amelynek működése egy fizikailag létező robot irányítása révén nem lép ki a fizikai világba, hanem pusztán szoftveres keretek között marad. Utóbbira példa lehet a már említett sakkprogram, vagy bármely más számítógépes játék.

RUSSEL és NORVIG meghatározása szerint a robotok olyan fizikai ágensek, amelyek a fizikai világ megváltoztatásával oldanak meg feladatokat. E célból különböző beavatkozó szervekkel szerelik fel őket, például lábakkal, kerekekkel, karokkal és megfogókkal. A beavatkozó szervek célja, hogy fizikai hatást fejtsenek ki a környezetre. A robotokat érzékelőkkel is felszerelik, hogy érzékelhessék a környezetüket (pl. kamera, ultrahangradar, giroszkóp, gyorsulásmérő). A robotok között megkülönböztethetünk a munka helyén rögzített robotokat (ún. manipulátorok) és helyváltoztatásra képes mobil robotokat. A robotika tárgyköre magában foglalja a protéziseket (pl. emberek számára készített mesterséges végtagok, illetve más szervek), az intelligens környezeteket (pl. egy egész ház felszerelve szenzorokkal és beavatkozó szervekkel) és a többemű rendszereket is, ahol a feladatokat sok apró együttműködő robot hajtja végre.²⁸

3.3 A mesterséges intelligencia jogi szabályozásának alapjai

A mesterséges intelligencia és a jog kapcsolatából fakadó problémák nem újkeletűek, azok immár több évtizede foglalkoztatják a tudományos élet művelőit. Az első utalást a szintetikus lények viselkedésének törvényi szabályozására azonban szépirodalomban találjuk.

ISAAC ASIMOV 1950-ben megjelent, kilenc történetet tartalmazó tudományos-fantasztikus novelláskötetében, az *Én, a robotban* írja le a robotika három alaptörvényét, a *Körbe-körbe* című elbeszélés során.²⁹ Asimov három alapszabálya, amelyet egy mesterséges lénynek feltétlenül követnie kell működése során a következő:

1. *A robotnak nem szabad kárt okoznia emberi lényben, vagy tétlenül túrnie, hogy emberi lény bármilyen kárt szenvedjen.*
2. *A robot engedelmessé válik az emberi lények utasításainak, kivéve, ha ezek az utasítások az első törvény előírásaiba ütköznek.*
3. *A robot tartozik saját védelméről gondoskodni, amennyiben ez nem ütközik az első vagy második törvény bármelyikének előírásaiba.*

Asimov később a fenti három törvényt kiegészítette a nulladik törvénnyel, amely szerint: „*A robotnak nem szabad kárt okoznia az emberiségben, vagy tétlenül túrnie, hogy az emberiség bármilyen kárt szenvedjen.*” Az eredeti törvényeket pedig kiegészítette a nulladik megszegésének tilalmával.³⁰ Mint látjuk ez az előírás a későbbi háromhoz képest elsődlegességet élvez, tehát a robot az első törvényt is megszegheti, ha ezzel az emberiség javát szolgálja.

A klasszikus ASIMOV féle törvények gyakorlati alkalmazásuk során azonban könnyen ellentmondásba keveredhetnek egymással. Mi történik például akkor, ha egy ember arra utasít egy robotot, hogy sértsen meg egy másik embert, mivel ez annak a javát szolgálja. Ez akkor fordulhat elő, ha például egy robot orvosi szolgálatban vesz részt és egy műtét végrehajtása során kap különböző utasításokat.³¹ Ezt az ellentmondást próbálta meg feloldani az író a nulladik törvény megalkotásával, amely arra utasítja a robotot, hogy működése során elsősorban az „emberiség mindenek felett álló érdekét” kell szem előtt tartania. Magunk is beláthatjuk azonban, hogy annak felmérése, hogy egy adott szituációban mi számíthat az emberiség érdekének rendkívül bonyolult és fejlett intelligenciát, továbbá ezen felül empátiás és morális képességét követel meg egy mesterséges tudattól. Nem csoda, hogy a törvényeit írói karrierje során maga ASIMOV is több tucatszor átírta, pontosította. Előfordult az is, hogy egyes történeteiben olyan robotokról mesélt, amelyekből szándékosan iktatták ki tervezőik valamelyik törvényt, hogy feladatuknak maradéktalanul eleget tudjanak tenni.³²

Igaz, ezek a törvények egyelőre megmaradtak a tudományos-fantasztikus irodalom szintjén, azonban MI fejlesztéssel a gyakorlatban is foglalkozó szakemberek nem vetik el annak a lehetőségét, hogy hasonló korlátozó tényezőkkel kellene felruházni a szintetikus létformákat.³³

ASIMOV törvényei azért is érdekesek, mivel először vetették fel annak lehetőségét, hogy a mesterséges intelligenciák viselkedését nem csupán technikai, hanem egy felsőbb szinten: a jog és a törvényesség szintjén is kívánatos lehet szabályozni. Asimov törvényeit felfoghatjuk egy olyan logikai rendszer elemeiként is, amelyek a robotot irányító szoftver szempontjából a szükséges és elégséges követelményeket jelentik.

Természetesen nem Asimov és science-fiction irodalma az egyetlen, amely foglalkozott a mesterséges intelligencia viselkedésének jogi szabályozhatóságával, hanem erre a jogtudományi szakirodalomban is találhatunk példát.

PETER M. ASARO a mesterséges intelligencia és robotika jogi vetületét veti fel és abból a feltevésekből indul ki, hogy első sorban azt kell megvizsgálnunk, hogy a hatályos törvények alkalmazhatóak-e az MI által generált problémákra. Ennek modellezésére a felelősség és kárveszélyviselés szabályait veszi alapul. ASARO a robotokkal kapcsolatos jogszabályok vizsgálatánál arra a következtetésre jut, hogy a robotra, mint kereskedelmi forgalomba hozott termékre minden további nélkül alkalmazandóak a termékfelelősség szabályai. Később a robotokról áttér az szoftverágensnek révén a megbízás, illetve a szerződéses viszonyok problémájára, amely alapján viszont véleménye szerint az ágens felhasználóját terheli a felelősség.³⁴

GIOVANNI SARTOR a szoftverágenset képviselőre alkalmas entitásoknak tekinti, álláspontja változásakor pedig abból indul ki, hogy manapság rengeteg szerződést kötnek automatikus szoftverek segítségével emberi beavatkozás, vagy felülvizsgálat nélkül. Ezek autonóm MI szoftverek. Sartor szerint a képviselő kapcsán szükséges, hogy kizárjuk a tudatosság és ez alapján a felelősség alkalmazhatóságát az ágens alkalmazása során. Mivel az ágensnek nincs tudata, ezért részére a viselkedése nem lehet felelősségteljes. SARTOR aprólékos jogfilozófia logikai levezetésében először feltárja, hogy a mesterséges intelligencia viselkedése hasonló az élő organizmusokéhoz, mivel azt előre megjósolni nem tudjuk pusztán belső működési elvei feltárásával, a sok véletlenszerű változó miatt. SARTOR hipotézise szerint azonban feltételeznünk kell, hogy a szoftverágens racionálisan viselkedik a számára meghatározott keretek és célok között (racionálisan cselekvő rendszert képez). Az ágens működéséért való felelősség nem azért terheli annak felhasználóját, mert kívánta, vagy előre látta annak viselkedését és következményeit, hanem azért mert eszközként kiválasztotta arra, hogy az által eredményeket érjen el, amelyeket elfogad, vagy akár fel is használ és ez számára jogokat és kötelezettséget keletkeztethet.³⁵

4. A mesterséges intelligencia, mint jogi entitás

4.1 Az MI, mint szoftver

Mivel a mesterséges intelligenciák különös jogi szabályozásának kísérletére eddig nem volt számottevőbb példa, ezért az előzőekben a filozófia és a technika fogalmait hívtuk segítségül annak definiálására, hogy mit is tekinthetünk annak.

A mesterséges intelligenciák viselkedése a jog szempontjából kétségteljesen akkor ér el egy kritikus pontot, amikor interakcióra kerül sor közte és az emberek, vagy tárgyak között. Erre legtöbbször akkor kerül sor, ha az MI révén működő szoftver valamilyen változást idéz elő az ember által érzékelhető környezetben, így például hatással van a környezetére, felhasználójára, illetve annak vagyontárgyaira. A kérdés ezzel kapcsolatban, hogy a szintetikus létforma cselekedeteiért kinek kell vajon vállalnia a felelősséget?

Ha a hatályos magyar jog alapján szeretnénk besorolni a mesterséges intelligenciákat, akkor először is azt kell szem előtt tartanunk, hogy azok nem mások, mint számítógépi programok, szoftverek.

A szoftver fogalmi behatárolására több megközelítés létezik a szakirodalomban. A fogalom körébe tartozhatnak többek között az utasítások, az adatstruktúrák és dokumentumok, érthető alatta a számítógép összehangolt működése is, valamint a hardverektől különböző, az adatfeldolgozó berendezések működtetéséhez szükséges szellemi terméket is jelentheti, amely az egyes programok gyűjtőneveként szolgál. A szoftver informatikai szempontból számítógépes programok, folyamatok és esetlegesen a számítógépes rendszer üzemelésére vonatkozó dokumentációk és adatok összességét jelenti. Más szemszögből a szoftver algoritmusokból és azok számítógépes reprezentációból, programokból áll. A szoftver csak akkor alkalmazható feladatok megoldására, ha az bizonyos

műveletek lépésenkénti végrehajtását eredményezi. Az algoritmus egy feladat elvégzéséhez szükséges, meghatározott lépések sora, amely bemenetenként bizonyos értéket, vagy értékeket kap és kimenetenként bizonyos értéket, vagy értékeket állít elő. Informatikai, mérnöki szemszögből nem alakult ki egységes szoftverfogalom, azonban a többségi vélemények alapján az legalább két részből áll: a számítógépes programból és a dokumentációból.³⁶

A magyar jogi szakirodalomban is számos álláspont létezik a szoftver fogalmának meghatározására. A szerzői joggal foglalkozó szerzők leggyakrabban a Berni Unió Egyezményre, a TRIPS Egyezményre, a WIPO Szerzői Jogi Szerződésre, valamint az EU szoftver irányelvére utalnak.³⁷

PÁLOS GYÖRGY a szoftvert három fő részre osztja, amelyek a forrásprogram (kód), a gépi/tárgyi program (kód) és a kísérő anyag (dokumentáció).³⁸

TATTAY LEVENTE szerint a számítógépes programalkotás és az ezzel kapcsolatos dokumentáció adja a szoftvert, amely szerinte irodalmi műként értékelhető. TATTAY a szoftvertől elvárja, hogy az a számítógép részeivel kapcsolatba lépjen, a felhasználóval kapcsolatot teremtsen, és az előre kitűzött eredményt elérje.³⁹

A LONTAI ENDRE, FALUDI GÁBOR, GYERTYÁNFY PÉTER és VÉKÁS GUSZTÁV féle egyetemi tankönyv szerint a szoftver megállapítások és utasítások sorozata, amely a számítógép vonatkozásában meghatározott eredményre vezet. Pálossal ellentétben a forráskódot és a tárgykódot nem a szoftver részeként, hanem a számítógépi programok rögzítési formáiként definiálja.⁴⁰

BALOGH ZSOLT GYÖRGY azon a véleményen van, hogy a számítógépi programalkotás nem irodalmi mű, hanem önálló alkotói műfaj, az egyéni eredeti jelleg speciális megjelenési formáival.⁴¹ DUDÁS ÁGNES egyetért ezzel a véleménnyel és szintén nem irodalmi műként, hanem sui generis alkotásként minősíti a szoftvert.⁴²

A fentiek alapján látszik, hogy eltérő álláspontok alakultak ki arról, hogy helyes-e az irodalmi besorolás, másrészt az, hogy nem félrevezetők-e a szerzői jog védelmének határai.

A hatályos magyar jogszabályok sem részletezik a szoftver fogalmát különösképpen, azonban a szerzői jogi törvény például külön is kitér arra, hogy a szoftver (a törvény alapján a számítógépi programalkotás és a hozzá tartozó dokumentáció) szerzői jogvédelem alatt álló alkotás (lásd: Sztj. 1. § (2) bekezdés c) pontja). A MI viselkedéséért való felelősség azonban nem (csak) szerzői jogi probléma, erről azonban a későbbiekben esik majd szó.

Fontos megjegyezni, hogy adott körülmények között akár a szoftver egy dolog alkotórészévé is válhat (pl. autó elektronikája). Ilyenkor a hordozó és benne a beégetett program osztja a dolog jogi sorsát. Ettől még a szerzői jogi védelem a szoftveren fennáll, de olyan módon kell a jogszerzést rendezni, hogy az alkotórészi „jogállás” megmaradhasson. Ez az eset a téma szempontjából az MI által irányított robotok esetében fordulhat elő.

4.2 Eddigi kísérletek az MI jogi fogalmának meghatározására

A szoftver fogalmához hasonlóan az MI fogalmának meghatározására is kevés kísérlet született eddig a világon. Erre példa az USA New Jersey államában 2014-ben elfogadott törvény, amely a vezető nélküli gépkocsikra vonatkozó különös előírásokkal foglalkozik és többek között előírja, hogy ezek üzemeltetői különleges engedéllyel kell, hogy rendelkezzenek.⁴³ A törvény meghatározása szerint mesterséges intelligenciáról akkor beszélünk, ha a számítógépet és azzal összekapcsolt berendezéseket olyan célra használják, hogy egy gép utánozza, vagy lemásolja az emberi viselkedést.⁴⁴

Véleményem szerint ez a kísérlet általános meghatározásnak elégtelen az MI fogalmának definiálása szempontjából, mivel az alatt kizárólag az emberi viselkedés modellezésére írt MI szoftvereket érti és ezek közül is csak azokat, amelyek valamilyen hardveres közeget (pl. egy robotot) működtetnek. Fentebb már többször ki lett emelve, hogy az MI fejlesztés célja általában nem az emberi viselkedés utánzása, hanem racionálisan cselekvő rendszerek építése, továbbá nem csak gépek és robotok irányítására írnak ilyen szoftvereket, hanem attól teljesen eltérő célokra is.

A hatályos magyar törvények nem tesznek különbséget a számítógépi programok szabályozása között a tekintetben, hogy azok mennyire rendelkeznek önálló döntéshozatali mechanizmusokkal és nem határozzák meg a mesterséges intelligencia fogalmát sem. Ez nem feltétlenül probléma, hiszen a felmerülő kérdéseket a jelenleg hatályos szabályozás alapján is legtöbb esetben meg lehet oldani, ahogy az alábbiakból kiderül. A jövőben az MI-vel foglalkozó esetleges különös törvényi

szabályozásban azonban vélhetően meg kell majd határozni annak fogalmát, a tárgyi hatály kijelölése céljából.

4.3 A *de lege ferenda* álláspont: MI mint jogalany

Az alábbiakban az MI jogalanyisággal való felruházhatóságával foglalkozó elméletek bemutatásával vizsgáljuk meg, hogy a mesterséges intelligenciát, olyan önálló jogi léttel bíró entitásnak tekinthetjük-e. Ezzel kapcsolatban a megválaszolendő fő kérdések, hogy illethetik-e jogok, illetve terhelhetik-e kötelezettségek önmagában az MI-t, vagy ha nem, akkor lehet-e azon keresztül érvényes jognyilatkozatot tenni.

A mesterséges intelligencia jogalanyisággal való felruházásának lehetősége már az 1990-es évek eleje óta heves vita tárgyát képezi a szakirodalomban.⁴⁵ FRANCISCO ANDRADE és szerzőtársai 2007-es tanulmánya rámutat a problémára, hogy a számítástechnika, a telekommunikáció és a mesterséges intelligencia fejlődése a szerződések megkötésének, továbbá a szerződési akarat kifejezésének új módjait hozta létre a közelmúltban. A vállalati szférában az intelligens elektronikus ágensek napjainkban egyre inkább jelen vannak, amik működésüket tekintve olyan szoftverek, melyek megbízóik nevében fejtenek ki joghatást kiváltó tevékenységeket közvetlen emberi irányítástól mentesen.⁴⁶

Az ágens kifejezés a latin *agere* szóból származik, ami annyit jelent, hogy cselekedni. Általános definíciójuk szerint ágensnek nevezzük az olyan mesterséges kreatúrákat, amelyek programozható voltak miatt jól használhatók szimulációra, modellezésre és egyéb szabályozó mechanizmusok tanulmányozására. A létrehozott ágens vagy „test” lehet önálló (autonóm ágens), azt behelyezhetik egy környezetbe, ezáltal a *lény* az adott környezet ingereinek hatására bekövetkező viselkedésében fog megnyilvánulni. A test és a környezet lehet számítógépes szimuláció eredménye, de lehet valódi környezetbe helyezett robot is, mely például egy hálózat irányítása alatt áll. Lényeges vonása ennek a módszernek, hogy az ágens nem csak feldolgozza a beérkező jeleket, hanem – mivel „teste” van – vissza is tud hatni a környezetére, és szenzorjai révén folyamatosan érzékeli saját működésének eredményét, ami befolyásolni fogja további viselkedését. Így tehát egy folyamatos visszacsatolás valósulhat meg, akárcsak az élő rendszerek esetében. Az ágensek így az életjelenségek utánozására kitalált műszaki-szellemi eredmények. Az elektronikus ágens, egy adott feladatot valaki más nevében, a felhasználó közbeavatkozása nélkül ellátó szoftver, amely kommunikál más ágensekkel, más folyamatokkal és a környezetével is.⁴⁷

Az elektronikus ágensek definícióját törvényi szintre emelte az amerikai Uniform Electronic Transactions Act (UETA) 2. § (6) bekezdése, amely szerint ennek kell tekinteni minden számítógépes programot, illetve elektronikus, vagy más automatizált eszközt, amely más személy irányítása vagy felügyelete nélkül önállóan cselekszik, illetve reagál elektronikus jelzésre egészben, vagy részben.⁴⁸

A fentiek alapján elmondható, hogy gyakorlatilag az ember helyett, önállóan cselekvő entításokkal van dolgunk, amelyek azonban bizonyos, az ember által meghatározott tágabb, vagy szűkebb célokat szolgálnak.

A kérdés, hogy vajon a polgári jog keretei között lehet-e lehetőségünk arra, hogy az önállóan cselekvő mesterséges entításokat jogalanyisággal ruházzuk fel? A kérdés megválaszolásra először tekintsük át, hogy a polgári jog jogi személy meghatározásnak megfeleltethető-e az MI.

4.4 A jogi személy fogalmának történeti meghatározása

A jogi személy fogalmának első meghatározásakor, a XIII. század derekán, azokat személyösszességekre (*universitas personarum*) és vagyonösszességekre (*universitas rerum*) csoportosították a kánonjogban. Kedvezett ezen elmélet kidolgozásának a korabeli gazdasági-politikai helyzet, amikor egyrészt szükség volt az egyházi vagyon jogi rendezésére, másrészt sorra alakították a testületeket, céheket, önkormányzatokat. A kánonjogászok III. INCE PÁPA (1243–1254) vezetésével kidolgozták az ún. „fikciós elméletet” amely ezután minden későbbi elmélet alapja lett. Eszerint a jogi személy olyan szervezet, amely nem jogalany, de úgy kell tekinteni, mintha az lenne. Később az osztrák jogirodalomban GEORG ARNOLD HEISE volt az első, aki *Grundriss eines Systems des gemeinen Zivilrechts* (1807) című művében figyelmeztetett arra, hogy a jogi személy fogalmát nem szabad csak

a személyegyesülésekre korlátozni. Ehhez kapcsolták azt a gondolatmenetet, hogy a megengedett egyesületnek ugyanolyan jogképessége van, mint az embereknek. A jogi személy fogalma *Kant* filozófiájának hatása alatt első sorban SAVIGNY munkásságának eredményeként alakult ki. Maga a jogi személy elnevezés is az ő tanítása alapján terjedt el. A szervezetek jogalanyiságát biztosító jogi személy elmélete is kezdetben az ember jogalanyiságát vetítette rá a szervezetekre. A XX. századra a gazdasági fejlődés révén a fikciós elmélet következetes alkalmazása egyre nehezebbé vált. Egyre többször fordult elő, hogy ugyanaz a személy többféle vagyonnak állt az élén, vagy több ember egy bizonyos cél érdekében egyesült egymással. A társasági vagyon ilyenkor is elkülönült a tagok magánvagyonától. ALOIS BRINZ német jogtudós ezért fejlesztette ki a célvagyon elméletét *Lehrbuch der Pandekten* (1857) című művében, amely minőségi változást eredményezett a fikciós elmülethez képest. Szerinte kétféle vagyon létezik: olyan, amely egy természetes személy tulajdonában van (*Personenvermögen*) és olyan, amelyet valamely cél érdekében állítottak (*Zweckvermögen*). Elsősorban az alapítvány jellegzetességei alapján kifejtette, hogy a jogi személy esetén a vagyon az egyes emberektől különült el. Adott cél szolgálatában álltak a jogok és a vagyonok. Nem azt kell kérdezni, hogy kié a vagyon, hanem hogy milyen célt szolgál. Végül mindezeknek az elméleteknek az egybefoglalásával alkotta meg MOÓR GYULA Kelsen tiszta jogtanához kapcsolva saját elméletét, amely szerint a jogi személy olyan jogi szervezet, amelynél a jog bizonyos emberek cselekményeit nem saját maguknak, hanem a tőlük különálló szervezetnek számítja be (ún. beszámítási elmélet).⁴⁹ A jogi személy fogalmának fejlődése kapcsán így megállapítható, hogy az olyan társadalmi szervezeteket jelent, amelyek elkülönült vagyonnal rendelkeznek és jogképességük révén saját nevükben – képviselőjük útján – jogokat szerezhetnek és kötelezettségeket vállalhatnak.

4.5 A szoftverágens jogi személyiséggel való felruházásának korlátai

A hatályos magyar szabályozás nem tekinti külön jogalanyisággal bíró entitásoknak – így különösen jogi személynek – a mesterséges intelligenciákat. Ennek ellenére a szoftverágensok gazdag tudásbázissal rendelkeznek, továbbá bonyolult beállításai lehetővé teszik, hogy feladatokat tervezzenek, reagáljanak, tanuljanak, valamint kommunikáljanak és együttműködjenek más MI szoftverekkel. Önálló léttel ugyan nem bírnak, viszont önállóan járnak el és teljesítenek feladatokat bizonyos szituációkban. ALLAN és WIDDISON szerint a számítógépes szoftverek elérték azt a szintet napjainkra, hogy nem csupán automatikusan, hanem önállóan cselekszenek. A szoftverek szerepe a kereskedelemben mindinkább áttevődik a passzív, technikai segítő oldaláról az aktív, önállóan cselekvő résztvevő oldalára.⁵⁰ BALL és BREESE szerint egyenesen azt kell észrevennünk, hogy egyes programok úgy vannak megalkotva, hogy akár érzelmeket is tudjanak imitálni, sőt önálló személyiségjegyeket mutassanak.⁵¹

Az elektronikus ágensekről elmondható, hogy a klasszikus jogi személyektől (pl.: alapítványok, gazdasági társaságok) az is megkülönbözteti őket, hogy nem szorulnak természetes személyek képviseletére. A jogi személynek önálló, a természetes személytől elkülönülő érdekei vannak. Az elektronikus ágensek nem szorulnak képviseletre, hanem éppen ellenkezőleg, az önálló döntéseikkel képviselnek egy harmadik személyt.

Ahhoz, hogy egy MI-t jogi személynek tekinthessünk, elkülönült vagyonra és vagyoni érdekekre lenne szüksége, mivel a személynek magánjogi oldalról nézve elsősorban, mint a vagyoni forgalomban önállóan részt vevő entitásnak van relevanciája. Az MI tehát nem tekinthető jogi személynek, ezen álláspont felülbírálata pedig a technika jelen állása szerint nem is tűnik indokoltnak, mivel az sem önálló tudattal, sem elkülönült vagyonnal nem rendelkezhet.

4.6 Az MI, mint képviselő: az UNCITRAL modell törvény lehetséges megoldása

Ennek ellenére annyi azonban mindenképp elmondható, hogy az MI révén emberi beavatkozás nélkül lehet érvényes jognyilatkozatot tenni, ami azonban az általa képviselt személyt kötelezi. Ez összhangban van az UNCITRAL elektronikus kereskedelmi modell törvénnyel is, amelynek rendelkezései szerint az olyan üzenetek, amelyeket közvetlen emberi beavatkozás nélkül generálják az információs rendszerek, nem lehetnek alkalmasak arra, hogy közvetlenül a rendszert ruházzák fel

jogokkal és kötelezettségekkel. Ebben az esetben úgy kell tekinteni, hogy az ilyen „automatikus szerződési nyilatkozat” attól a jogi entitástól származik, aki a rendszert üzemelteti.⁵² Az ún. „automatikus szerződési nyilatkozat” azonban nem tévesztendő össze a képviseléssel. A képviselés annyit jelent, hogy egy jogalany – akár természetes, akár jogi személy – nevében valamely más személy tesz jognyilatkozatot úgy, hogy a jognyilatkozat joghatásai közvetlenül annál a személynél, azaz a képviselnél állnak be, akinek a nevében a nyilatkozatot megtették. A Ptk. alapján képviselő csak személy lehet, ágens nem. Ennek alapja, hogy adott esetben a képviselőnek elkülönült vagyoni felelőssége lehet, például, ha a képviselt érdekének ellenében cselekszik, és ezzel annak kárt okoz. Érdekellentét esetén a képviselt a képviselő jognyilatkozatát akár meg is támadhatja, amely alatt a szerződések érvénytelensége körében szabályozott megtámadást (Ptk. 6:89. §) kell érteni.⁵³ Ez a megtámadási lehetőség értelemszerűen egy szoftver által tett jognyilatkozat kapcsán nem fordulhat elő.

Az MI ágens tehát a hatályos jog alapján nem tekinthető se jogi személynek, se képviselőnek, viszont azon keresztül jognyilatkozatot tenni lehet, annak ellenére, hogy a magyar jog az automatikusan szerződési nyilatkozat fogalmát nem kifejezetten ismeri. Emiatt véleményem szerint szerencsés lenne az UNCITRAL fogalmának implementálása a magyar polgári jogba.

A fentiek alapján tehát sikerült behatárolni, hogy az MI jogi fogalmának kidolgozása még várat magára, így jelenleg elsősorban az kell szem előtt tartanunk, hogy azok szoftverek, vagy adott szoftver részeként működő algoritmusok. Az MI önálló jogi személyiséggel történő felruházhatósága mellett kevés érv szól, így az alábbiakban a felelősségi rendszer felvázolásakor első sorban abból kell kiindulni, hogy egy autonóm módon működő szoftver viselkedéséért kinek kell viselnie a felelősséget.

5. A mesterséges intelligencia viselkedéséért való magánjogi felelősség a szerződéses viszonyokban

Vajon a polgári jogi kárfelelősség szabályai szerint mi történik akkor, ha egy mesterséges intelligencia valamilyen kárt okoz működése során? Mint az előzőekben kifejtettek alapján láttuk, a mesterséges intelligencia egy szoftver, tehát első sorban azt kell megvizsgálnunk, hogy a szoftverek működéséből eredő káresemények bekövetkezéséért kiknek kell vállalniuk a felelősséget. Ebben a pontban a szerződéses viszonyok keretében beállott károkért való felelősség problémájának a bemutatása következik.

5.1 Jogirodalmi álláspont

A jogirodalomban kevés álláspont létezik a fenti problémával kapcsolatban. A már idézett GIOVANNI SARTOR egy tanulmányában kitér a szoftverágens által okozott károkért való felelősség problémájára. Először számba veszi, hogy egy bizonyos szoftverágens elkülönült jogi sorsot osztó „részei” révén több jogalanyal is számolni lehet a felelősségviselés telepítésének szempontjából. Ha az ágens szerzői jogvédelem alatt álló szoftvert tartalmaz (mint legtöbb esetben), akkor a szerző viselheti a felelősséget a programozás során vétett hibák miatt. Ha az ágens valamilyen adatbázist tartalmaz, akkor az ebből fakadó hibákért az adatbázis előállítója lesz a felelős. Ha az ágens személyes adatokat kezel, akkor adatvédelmi szempontból az adatkezelő felelős a személyes adatok kezelésének jogszerűségéért. Ha az ágenst egy adott felhasználó üzemelteti saját céljaira, akkor a működtetéséért a felelősség is őt terheli.⁵⁴

Mivel a szoftverágens használatának vannak olyan aspektusai, amely felett adott jogalany nem képes felügyeletet gyakorolni, ezért az e szempontból okozott kárért nem terhelheti felelősség Sartor szerint: Például nem tehető felelőssé a felhasználó a szoftver programozási hiányosságából eredő működéssel összefüggésben okozott károkért, ha nincs hozzáférése a szerző által írt forráskódhoz, ráadásul annak visszafejtésére sem jogosult.

5.2 A felhasználó felelőssége

Ki viseli tehát az MI szoftverek működéséért a felelősséget? Először is azt kell szem előtt tartanunk, hogy a számítógépes programok szerzői művek. A szerzői művek értékesítésére, felhasználására a szerzői jogi törvényben meghatározott szabályok irányadóak.

Amikor egy felhasználó megvásárol egy szoftvert a piacon, jellemzően meg kell kötnie a felhasználási szerződést, általában a szoftver használatával kapcsolatos általános szerződési feltételek elfogadása útján. A szerződés módosítására a felhasználónak legtöbbször nincs lehetősége, így csak arról dönthet, hogy elfogadja-e az abban foglaltakat, vagy sem.

A szerzői jogi törvény definíciója szerint a szerző az ilyen felhasználási szerződésben engedélyt adhat művének a felhasználására, a felhasználó pedig – ha a felek eltérően nem rendelkeznek – köteles ennek fejében díjat fizetni. A szerződésre vonatkozó rendelkezésektől egyező akarattal eltérhetnek a felek, feltéve, ha azt törvény, vagy más jogszabály nem tiltja.⁵⁵ A fentiek alapján tehát a kárfelelősség alapját a szerződéses viszonyokban okozott károk szabályozása között kell keresnünk, hiszen a szerző és a felhasználó között tipikusan szerződéses jogviszony jön létre a szoftver használatának körülményeit rögzítő felhasználási szerződés felhasználó általi elfogadásával.

A szoftverek forgalmazói az általuk értékesített programhoz mellékelt általános szerződési feltételeket tartalmazó dokumentumba kifejezetten bele szokták írni, hogy a felhasználó által önmagának, vagy harmadik személynek a szoftver használatával összefüggésben okozott károkért nem a szerző (vagy forgalmazó), hanem a felhasználó felel.⁵⁶ Ezt angol kifejezéssel élve „limited liability” (korlátozott felelősség) kikötésnek nevezik az egyes szoftverekhez mellékelt szerződési feltételekben.

5.3 A hibás teljesítés esete

Ettől meg kell különböztetni azt az esetkört, ha a szoftver hibásan került a piacra és e hibájából eredően álltak be károk a felhasználói oldalon. A Ptk. a hibás teljesítés jogintézményén keresztül enged lehetőséget arra, hogy a szolgáltatás nyújtóját felelősségre lehessen vonni a szerződés hibás teljesítése, jelen esetben egy hibásan megírt számítógépes program szolgáltatása miatt. A hibás teljesítés a szerződésszegés egyik nevesített esete.

A Ptk. 6:157 § (1) bekezdés alapján „*a kötelezett hibásan teljesít, ha a szolgáltatás a teljesítés időpontjában nem felel meg a szerződésben vagy jogszabályban megállapított minőségi követelményeknek. Nem teljesít hibásan a kötelezett, ha a jogosult a hibát a szerződéskötés időpontjában ismerte, vagy a hibát a szerződéskötés időpontjában ismernie kellett.*”

Ez a fogalom meghatározás irányadó mindazon szerződések hibás teljesítésének megítélése során, amelyeknél a szerződésszegés e típusa a szerződéses szolgáltatás természetére figyelemmel szóba jöhet, így a használati kötelektnél is.⁵⁷ A felhasználási szerződés tipikus használati kötelek, így a hibás teljesítésből eredő jogok érvényesítése kiterjedhet a szerződés kapcsán szolgáltatott szoftver hibájára is.

Ha a szolgáltatás, így jelen esetben a szoftver minőségével kapcsolatban sem jogszabály, sem a felek szerződése nem fogalmaz meg különösebb elvárásokat, akkor e vonatkozásban a Ptk. 6:123. §-ának (1) bekezdésében megfogalmazott általános követelmények jönnek figyelembe. Eszerint a szolgáltatásnak alkalmasnak kell lennie a rendeltetése szerinti, tehát az ilyen szolgáltatások esetében általánosan megszokott céljának betöltésére, és rendelkeznie kell az azonos fajtájú szolgáltatások szokásos, elvárható minőségével. A szolgáltatás tehát mindezekre figyelemmel hibás, ha bármely okból nem alkalmas, a rendeltetésének megfelelő felhasználásra, vagy nem rendelkezik az ilyen szolgáltatások szokásos, a jogosult által elvárható minőségével.⁵⁸

A hibás teljesítéshez kártérítési igény is tartozik, amely annak további jogkövetkezménye. Ez alapján a kötelezett köteles megtéríteni a jogosultnak nyújtott szerződés hibás teljesítéséből eredő kárát.⁵⁹

A hibás teljesítés folytán a jogosultnak többféle kára keletkezhet. A kár egyrészt jelentkezhet abban, hogy maga a szerződéses szolgáltatás hibás (ún. tapadó kár), másrészt a hibás teljesítés elindíthat egy olyan oksági folyamatot, amelynek következtében a jogosult egyéb vagyontárgyai károsodhatnak, illetve ennek elhárítása érdekében kiadásai keletkezhetnek (ún. következménykárok). A visszerhes szerződést hibásan teljesítő kötelezettnek a jogosult tapadó és következménykárát is meg kell térítenie.⁶⁰

A jogosult a kár megtérítéseként annak az értékcsökkenésnek a megtérítésére tarthat igényt, amelyet a hiba folytán elszenvedett, vagy a hiba kijavítási költségének megfizetését követelheti. Az általános igényérvényesítési idő egy év, a fogyasztó és vállalkozás között létrejött szerződés esetén két év.

Nem tartozik felelősséggel a kötelezett, ha a hibás teljesítést kimenti.⁶¹ Ez úgy fordulhat elő, ha a Ptk.-nak a szerződésszegés általános szabályai között elhelyezett 6:142. §-ában foglaltak szerint bizonyítja, hogy a hibás teljesítést ellenőrzési körén kívül eső, a szerződéskötés időpontjában előre nem látható körülmény okozta, és nem volt elvárható, hogy a körülményt elkerülje, vagy a kárt elhárítsa.⁶²

Ha a Ptk. fenti rendelkezéseit alkalmazzuk a szoftver forgalmazója és a felhasználó között létrejött jogviszonyra, megállapíthatjuk, hogy a felhasználó a hibásan megírt szoftver használatából bekövetkező károk megtérítését követelheti a hibás szoftver forgalmazójától.⁶³

5.4 Kellékszavatossági igények

Tovább árnyalja a képet, hogy a hibásan megírt szoftverért a szoftverfejlesztő kijavítási, „frissítési” felelősséggel is tartozhat a hibás teljesítéshez kapcsolódó kellékszavatosság jogintézménye alapján. A kellékszavatosság Ptk.-beli előírásai alapján a jogosult (a szoftvert futtató felhasználó) a szolgáltatás (tehát a felhasználóval kötött szerződés alapján futtatott szoftver) kijavítását, vagy kicserélését igényelheti, az ellenszolgáltatás arányos leszállítását igényelheti, a hibát a kötelezett költségére maga kijavíthatja vagy mással kijavíttathatja, vagy a szerződéstől akár el is állhat.⁶⁴ Ha a szerződés nem zárja ki, a hiba kijavítását a szoftver jogszerű felhasználója is elvégezheti.⁶⁵ A kellékszavatossági igények fő szabály szerint a teljesítéstől számított egy éven belül elévülnek, azonban használatra irányuló szerződéseknél a kötelezett a szerződés teljes tartama alatt a kellékszavatosság szabályainak megfelelő alkalmazásával köteles helytállni azért, hogy az oltalom tárgya (jelen esetben a szoftver) a szerződésszerű használatra, felhasználásra vagy hasznosításra alkalmas legyen.⁶⁶

A szoftverfejlesztők szinte mindig kiadnak az általuk kínált programokhoz különböző javításokat, ha saját észlelésük, vagy a felhasználók visszajelzése alapján annak működése során hibák merülnek fel. Ez érdekében is áll a szoftverek forgalmazóinak, hiszen egy hibásan működő program nem számíthat túl nagy érdeklődésre a piacon.

Fontos kihangsúlyozni, hogy a kártérítés szabályai diszpozitívak, azaz a felek azoktól a létrejött szerződésben eltérhetnek, tehát lehetséges a kártérítési felelősségnek a szerződésben történő korlátozása, vagy kizárása. Nem lehet azonban korlátozni vagy kizárni a szándékosan okozott, életet, testi épséget, vagy egészséget megkárosító szerződésszegésért való felelősséget.⁶⁷

5.5 Az MI szoftver működéséért való felelősség általános szabályai a szerződéses viszonyokban

Konklúzióként levonható, hogy az MI szoftver működéséért a szerződéses jogviszonyokban a felelősséget több jogalany is viselheti. Annak eldöntése kapcsán, hogy ez épp melyiknek a kötelessége (fejlesztő, forgalmazó, felhasználó stb.), mindig a probléma természetéből kell kiindulni, és, hogy az melyikük mulasztásával van okozati összefüggésben:

- A rendeltetésszerűen működő szoftver által a vagyonban beállott károkért a felhasználó felelős.
- A hibásan működő szoftver által a felhasználónak okozott kár megtérítését követelheti a felhasználó a fejlesztőtől (forgalmazótól) a hibás teljesítés szabályai alapján. Ez a tapadó károk teljes, a következménykárok és az elmaradt bevétel előreláthatóságához kötött megtérítéséhez vezethet.
- A hibásan forgalmazott szoftver kijavítását is követelheti a felhasználó annak fejlesztőjétől (forgalmazójától) a kellékszavatosság intézménye alapján.

Mivel azonban alapvetően diszpozitív szabályokról van szó, alapvetően a felek szerződése irányadó a kockázatmegosztásra.

A fenti rendszer azonban csak arra alkalmas, hogy a szerződéses jogviszonyban a másik szerződő fél (jogosult) oldalán bekövetkező kár megtérítésére kötelezze adott esetben az ellenszolgáltatás nyújtóját. Ha a felhasználó az MI szoftver futtatásával harmadik személynek, szerződésen kívül okoz kárt, arra más szabályok irányadók. Ezt az esetet az alábbi pontban tekintem át.

6. A szerződésen kívüli kárfelelősség általános szabálya: a mesterséges intelligencia, mint veszélyes üzem

6.1 Jogirodalmi álláspontok

A szerződésen kívül okozott károkért való felelősség elméleti és gyakorlati kérdéseivel számos magyar jogász foglalkozott a jogirodalomban. A magyar jogi gondolkodásra meghatározó jelentőségű volt MARTON GÉZA munkássága, a régi Ptk. szabályainak kialakítására pedig EÖRSI GYULA gyakorolt döntő befolyást. Az új, hatályban lévő Ptk. előkészítését végző szakértői bizottságnak a kártérítési felelősség szabályozására vonatkozó állásfoglalását LÁBADY TAMÁS alapozta meg.⁶⁸ A mesterséges intelligencia viselkedésének és a szerződésen kívüli kártérítési felelősség problematikájának feldolgozására a magyar szakirodalomban nem volt még példa, és nemzetközi viszonylatban sem találunk arra sokat. Az egyik ilyen a már idézett GIOVANNI SARTOR tanulmánya, aki érinti a probléma alapfeltevéseit.

SARTOR két elméletet állít fel a mesterséges intelligencia használatát illetően a kárveszélyviselés szempontjából. Az egyik, hogy a jogalany csak az olyan károkért tehető felelőssé, amelyek szempontjából minimum gondatlanság terheli, ezért mentesül a felelősség alól, ha az ágens magatartása nem volt neki felróható. Ennek az elméletnek a hátulütője, hogy a jogosultak sem képesek előre látni az MI ágens viselkedését, mivel az irányítási képességük határain túl is képes cselekedni, így a legtöbb esetben senki sem tehető majd felelőssé az okozott kárért.⁶⁹

A másik elmélet szerint az érintett mindig viseli a kárfelelősséget az ágens viselkedése kapcsán, függetlenül az elvárhatóság elvétől és így a felróhatóságtól. Ez a megközelítés a veszélyes üzemi kárfelelősség logikáját alkalmazza a mesterséges intelligenciára. A megoldás ésszerűnek tűnik, tekintettel arra, hogy az ágens viselkedése az önálló döntéshozatali mechanizmusok révén nem tervezhető előre, ezért annak üzemeltetése fokozott veszéllyel jár.⁷⁰

6.2 A veszélyes üzemi felelősség alapvető ismérvei

A fokozott veszéllyel járó tevékenység lényege, hogy csekély hiba is aránytalanul súlyos károkat eredményezhet. Sem a Ptk., se más jogszabály nem definiálja a veszélyes üzem fogalmát és nem is sorolja fel taxatíván az ilyen jellegű tevékenységeket, hiszen a technika és az elemi erők elleni védekezés fejlődésével az koronként változik.⁷¹ A jelenlegi gyakorlat szerint a gépjárműveken túl ide tartoznak többek között a gépi erővel működő gazdasági üzemek, a villamos energiát termelő és átalakító üzemek, továbbá a robbanó-, sugárzó-, mérgező anyagok, lőfegyverek készítése, tartása, használata is.

A veszélyes üzemi felelősség lényege, hogy annak fenntartója a károkért akkor is felel, ha őt, illetve megbízottját a kár bekövetkeztével összefüggő vétkesség nem terheli.⁷² Az általános felelősségi alakzatnál ez szigorúbb esetkör, ugyanis ha a kárt fokozott veszéllyel járó tevékenység folytatásával okozzák, akkor a felelősség nem a szubjektív alapú felróhatóságon alapul, hanem attól független objektív (tárgyi) felelősség. Ennek indoka, hogy az átlagshoz képest fokozott veszéllyel járó tevékenység folytatása az átlagost meghaladó védekezés megteremtését követeli meg, ha pedig a kár bekövetkezett, az azt jelenti, hogy a védekezés nem volt megfelelő.⁷³

A Ptk. meghatározza az üzembentartó fogalmát, amely szerint a fokozott veszéllyel járó tevékenység folytatójának az minősül, akinek érdekében a veszélyes üzem működik.⁷⁴ A Ptk. kommentárja szerint az üzembentartónak tehát azt kell tekinteni, aki a veszélyes üzemet fenntartja, tartósan üzemelteti és akinek fennáll a felügyelete, irányítása, ellenőrzése és ehhez kapcsolódik a különleges védekezésre való kötelezettsége.⁷⁵ Általában a tulajdonos üzembentartó is, ez a minőség azonban adott esetben el is válhat a tulajdontól.⁷⁶

A veszélyes üzem üzembentartója csak kivételes esetben mentesülhet a kárfelelősség alól, még pedig akkor, ha bizonyítja, hogy a kárt olyan elháríthatatlan ok idézte elő, amely a fokozott veszéllyel járó tevékenység körén kívül esik. Ezen egyetlen kimentési ok megvalósulásához a két feltétel együttes fennállta szükséges. Az elháríthatatlanság csak akkor állapítható meg, ha – a technika adott fejlettségi szintjére és a gazdaság teherbíró képességére is figyelemmel – objektíve nem áll fent a védekezés

lehetősége. Azt a körülményt, hogy objektíven fennáll-e a védekezés lehetősége, mindig a konkrét ügyben, az adott társadalmi viszonyok közepette kell értékelni, szem előtt tartva, hogy az egyensúly megteremtése mellett is az a végső cél, hogy a kár elhárítható legyen.⁷⁷

A megítélés szempontjából tehát azt kell figyelembe venni, hogy létezik-e olyan megoldás, amellyel a kár bekövetkezte elkerülhető lett volna. A károkozó nem mentesül a kártérítési felelősség alól annak bizonyításával, hogy a szabályoknak és a műszaki előírásoknak megfelelő megoldásokat alkalmazott, vagyis nem ezek betartásának van jelentősége, hanem annak, hogy a fokozott veszéllyel járó tevékenységet folytató ezeken túlmenően is mindent megtett-e a kár elhárítása érdekében.

A másik szükséges feltétel, hogy a felelősség alóli mentesülésre vezető elháríthatatlan oknak külső tényezőnek, a veszélyes üzem tevékenységi körén kívül eső körülménynek kell lennie. A bírói gyakorlat szerint nem mentesít a felelősség alól a fel nem ismerhető okból bekövetkező rendellenesség, amely magában a veszélyes üzemben lépett fel.⁷⁸ Ilyen ok lehet például az MI szoftver hibás, rendellenes működése is, ha azzal szerződésen kívül okoznak kárt.

6.3 A veszélyes üzemi felelősség modelljének alkalmazása a MI-t érintő viszonyokra

A fentiekben ismertetett előírások a mesterséges intelligenciát felhasználó fokozott veszéllyel járó tevékenységekre – elfogadva SARTOR véleményét – alkalmazhatóak. A szerződéses viszonyokon kívül okozott károkért tehát a szoftver üzemeltetői felelnek fő szabály szerint, akik legtöbbször a felhasználók. Az MI szoftver üzemeltetőjének tehát fő szabály szerint mindig viselnie kell a felelősséget, azt nem háríthatja át másra, mivel az MI viselkedése előre nem tervezhető. Kivétel lehet viszont, ha sikerül bizonyítani, hogy a szoftver működési körén kívüli külső tényezővel van okozati összefüggésben a káresemény bekövetkezése.

Az MI-nek azonban számtalan olyan felhasználási módja ismert, amely – ellentétben a klasszikus veszélyes üzemi tevékenységekkel – alkalmatlan arra, hogy kárt okozzon. Ilyenek például a számítógépes játékok szkriptek által irányított ellenfelei (lásd: sakkprogram működési elveinek korábbi leírását), illetve egyéb programozott eseményei (ún. eventek), egyes intelligens adatbázisok, vagy a virtuális valóság szoftverek felhasználói által létrehozott automatizmusok nagy része is.

Önmagában azonban ez véleményem szerint nem indokolja, hogy a kárfelelősség szempontjából különbséget tegyünk MI és MI között. Ha az MI alkalmas arra, hogy kárt okozzon (bármekkora is legyen az), akkor ésszerű, hogy fokozott veszéllyel járó tevékenységnek tekintsük üzemeltetését a fentiekben kifejtettek alapján, ha pedig károkozásra belső tulajdonságai miatt alkalmatlan, akkor a kárfelelősség kérdése fel sem merülhet annak üzemeltetésével kapcsolatban.

Mivel a mesterséges intelligenciák használatának térnyerése a gazdaságban előreláthatólag egyre nagyobb teret fog ölteni, érdemes lenne egyértelműen rendezni az azokkal kapcsolatban felmerült felelősségi kérdéseket. Ebből a szempontból követendő példa a már hivatkozott UNCITRAL modelltörvény, amely egyértelműen kimondja, hogy a mesterséges intelligencia üzemeltetés során, annak tevékenységéért és ezzel összefüggésben okozott kárért annak üzemeltetője felel. Az MI üzemeltetés ezen felül fokozott veszéllyel járó tevékenységnek minősül.

A mesterséges intelligenciáknak kivétel nélkül vannak üzemeltetőik, és azok az ő képviseletükben és általuk meghatározott célok érdekében működnek, valamint számukra szereznek jogokat és kötelezettségeket, így tetteikért a kárfelelősséget is nekik kell főszabály szerint viselniük.

7. A mesterséges intelligencia által okozott károkért való büntetőjogi felelősség

7.1 A büntetőjogi felelősség általános kérdései: az MI, mint elkövető?

GABRIEL HALLEVY 2010-ben megjelent tanulmányában veti fel a kérdést, hogy vajon mi a teendő, ha egy mesterséges intelligencia működése során olyan magatartást tanúsít, ami a büntetőjog szabályai szerint bűncselekmény törvényi tényállásának felel.⁷⁹

Fontos leszögezni, hogy a büntetőjogi felelősség megállapítására szűkebb keretek között kerülhet csak a sor a polgári kártérítési felelősséghez képest. Ahhoz egyrészt az elkövető által tanúsított magatartásnak

minden elemében ki kell mérítenie a Büntető Törvénykönyv valamelyik különös részi tényállásának konjunktív feltételrendszerét (*actus reus*), másrészt szükséges megvizsgálni az elkövetéskori tudattartamot is (*mens rea*). Ez azért fontos, mert a magyar Büntető Törvénykönyv fő szabályként csak a szándékos bűnelkövetést rendeli büntetni, tehát ha az elkövető tisztában van a cselekménye következményeivel és azokat kifejezetten kívánja, vagy legalábbis azokba belenyugszik.⁸⁰ A gondatlan elkövetőt csak kivételes, nevesített esetekben rendeli büntetni a törvény.

Az elkövető által megvalósított cselekmény fogalmával kapcsolatban több meghatározás létezik. A klasszikus iskola által képviselt kauzális cselekmény fogalom szerint a cselekmény olyan *akaratlagos* magatartás, amely a külvilágban meghatározott következményeket idéz elő. Ezzel szemben a finális cselekménytan hívei az okozatosságot kiegészítik a célirányozottsággal: szerintük a cselekmény nem csak az akarattól meghatározott kauzális folyamat, hanem lényege szerint célratoró tevékenység. Ezen elméleteket alapul véve a magyar szakirodalomban FÖLDVÁRI JÓZSEF a kettő elmélet vegyítésével a magatartás és gondolkodás, tudatosság (motívumok harca, melyet valamely elhatározás zár le) viszonyát emeli ki a büntetőjogi felelősség szempontjából.⁸¹

Ezek a jogi szabályok és általános elfogadott álláspontok már alapvetően rányomják a bélyegüket a mesterséges intelligencia által elkövetett tényállásszerű magatartásokra. A külföldi irodalomban UGO PAGALLO is osztja a nézetet miszerint a szoftverek – habár rendelkezhetnek bizonyos fokú döntéshozó mechanizmusokkal – a technika jelen állása szerint egyelőre nem rendelkeznek tudattal, ezért azok bűnös tudattartamának vizsgálata eleve lehetetlen, így az MI általi szándékos, illetve gondatlan bűnelkövetés lehetősége kizárt.⁸² Az MI-k nincsenek tisztában saját létezésükkel, illetve tetteik erkölcsi, etikai súlyával és következményeivel. Nem mérlegelik tetteik társadalmi súlyát, hanem egyszerűen csak azon keretek között cselekszenek, ahogy programozták őket.

A fentiek kifejtett érvek alapján eleve lehetetlen csupán a program bűnügyi felelősségre vonása bűncselekmények elkövetése esetén. Ez az állapot valószínűleg fent is fog állni, addig a pontig, amíg az intelligens viselkedést szimuláló szoftverek megmaradnak a *gyenge MI* szintjén, tehát tetteik valamilyen szinten kontrollálhatóak, vagy legalábbis azok keretei bizonyos szinten előre meghatározhatóak.

A hatályos törvények értelmében egy szoftver büntetőjogi felelősségét nem lehet vizsgálni, mivel a Btk. személyi hatálya kizárólag a természetes személyekre terjed ki,⁸³ továbbá – a korábban kifejtettek értelmében – jogi személyiséggel sem rendelkezik, így a jogi személyekkel kapcsolatos bűnügyi intézkedések sem terjednek ki rá.⁸⁴ A fentiek értelmében a mesterséges intelligencia működése önmagában jelenleg nem szankcionálható büntetőjogi eszközökkel.

Ennek ellenére mégis érdemes vizsgálni a problémát, hiszen számos olyan élethelyzet adódhat, ahol közvetve, vagy közvetlenül egy mesterséges entitás is érintve lehet bűncselekmények elkövetésében. Az alábbiakban az MI-t, mint elkövetési eszközt vizsgáljuk, hipotetikus példákön keresztül szemléltetve az elkövetési módokat.

7.2 Az MI, mint elkövetési eszköz

Előfordulhat, hogy egy mesterséges intelligencia ugyan tényállásszerű magatartást tanúsít, azonban azért teljes mértékben egy harmadik személy felelős. Példa erre, ha az elkövető egy MI szoftvert használ egy bűncselekmény elkövetéséhez.

Elkövetési eszköznek csak olyan eszköz minősülhet, amely az elkövető testétől függetlenül létezik.⁸⁵ Ez a meghatározás így az MI szoftverre is alkalmazható lehet.

Ha elfogadjuk, hogy a mesterséges intelligenciákat nem terheli közvetlenül büntetőjogi felelősség, mivel azok viselkedése mögött nem önnön tudatosságuk, hanem előre programozott parancsok húzódnak meg, akkor a parancsok kiadót tekinthetjük ilyenkor tetteseknek. Ha például egy gyári robotot úgy állítanak be, hogy munkaidő után menjen végig az épületen és gyújtson tüzet, aminek következtében a gyár leég, nem a robot, hanem az követi el a közveszély okozásának büntettét, aki erre utasította. Nyilvánvaló, hogy a robot önmagától, közvetlen irányítás nélkül fogja elkövetni a cselekményt, de azért mert erre kapott parancsot.

Hasonló helyzet állhat elő akkor, amikor a mesterséges intelligencia működésében valamilyen szándékos külső beavatkozás hatására áll be zavar. Tegyük fel például, hogy egy vezető nélküli autó informatikai rendszerébe illetéktelenek kívülről belépnek és átállítják, hogy az bizonyos sebesség felett kapcsolja ki a féket.⁸⁶ Ha a vezető emiatt balesetet szenved, az illetéktelen behatolót nem csak

testi sértés, vagy súlyosabb esetben emberölés büntetnének elkövetése miatt terheli a felelősség, hanem információs rendszer vagy adat megsértése bűncselekmény elkövetése miatt is, mivel egy számítógépes rendszerben a jogosulatlan adatmódosítás már önmagában is tényállásszerű.⁸⁷ A fenti esetekben a szoftver tehát nem más, mint a bűncselekmény elkövetésének eszköze.

8. Konklúziók: Az MI felelősségi rendszer felvázolása

Do you remember the question that caused the Creators to attack us, Tali'Zorah?: „Does this unit have a soul?”

Legion (Mass Effect 3)⁸⁸

A tanulmányban arra kerestük a választ, hogy a mesterséges intelligenciával ellátott szoftverek viselkedéséért és azzal összefüggésben okozott károkért ki tehető felelőssé. A technika fejlettsége alapján napjainkban a mesterséges intelligenciák a gyenge MI szintjén állnak, tehát azok olyan fokú önállóságról nem tesznek tanúságot, hogy tudatosan cselekvő létformáknak lehessen tekinteni őket. MI alatt a racionálisan cselekvő rendszereket értettem az elemzés során, mivel a gyakorlatban leginkább az ilyen viselkedést szimuláló szoftvereknek és szkripteknek van jelentősége és nem az emberi gondolkodás szimulálást megkísérlő társaiknak.

A hatályos jog vizsgálata alapján megállapítható, hogy a törvények nem szabályozzák külön a mesterséges intelligenciákra vonatkozó szabályokat, így azokat más általános jogi normákból tudjuk levezetni. A mesterséges intelligencia alkalmazása tehát képes arra, hogy előre programozott logikai módszerek révén önálló döntéseken alapuló és joghatás kiváltására is alkalmas tevékenységet végezzen közvetlen emberi behatás nélkül.

Ezek a technikai megoldások azonban sajnos a gyakorlati problémák szintjén jogi szempontból jó pár értelmezési problémát szülhetnek. Megállapítható, hogy a hatályos jogi normák nagyrészt alkalmasak arra, hogy a MI szoftver hibás, vagy éppen rendeltetészerű működéséből eredő felelősségi problémákat meg tudjuk oldani, a probléma újszerűsége és az egységes gyakorlat hiánya azonban megkíván egy hasonló bővebb, tudományos elemzést.

A gyenge mesterséges intelligencia viselkedéséért való felelősség rendszerét az alábbi négy pontban kísérlem meg felvázolni:

1. A mesterséges intelligencia (MI) üzemeltetése fokozott veszéllyel járó tevékenység.
2. Az MI viselkedéséért és azzal okozati összefüggésben keletkezett károkért annak üzemeltetője felel. Mentésül a felelősség alól, ha bizonyítja, hogy a kárt olyan elháríthatatlan ok idézte elő, amely a fokozott veszéllyel járó tevékenység körén kívül esik.
3. Amennyiben az MI bűncselekmény elkövetésének eszköze, úgy a büntetőjogi felelősség akkor állapítható meg, ha a szoftver magatartása tényállásszerű és azzal okozati összefüggésben az üzemeltető, vagy harmadik személy részéről a szándékos, vagy adott esetben gondatlan elkövetés ténye megállapítható.
4. A szerződéses viszonyokban az MI szoftver felhasználása kapcsán a hibás teljesítésből és kellékszavatosságból eredő jogok érvényesíthetőek. A kockázatot a felek szerződésükben eloszthatják.

Fontos kihangsúlyozni, hogy a fenti megállapítások csupán a gyenge MI szintjén tűnnek megalapozottnak. Ha a jövőben az emberiség eljut arra a szintre, hogy sikerül létrehozni az erős, önálló tudattal rendelkező MI-t, akkor a fejezetben lefektetett megállapítások sem fogják feltétlenül megállni a helyüket.

Addig is azonban szerencsésebb lenne úgy alakítani a törvényi szabályozást, hogy az ilyen különleges szituációkban is gond nélkül alkalmazható legyen. A jogalkotásnak rugalmasan időről időre reflektálnia kell az új társadalmi jelenségekre, ez pedig a csúcstechnológia fejlődése mentén különösen gyakran fordul elő.

-
- ¹Krausová, Alzbeta: *Legal Regulation of Artificial Beings*. Masaryk University Journal of Law and Technology, 2007. p. 188.
- ²Siba, László (szerk.): *Oxford számítástechnikai értelmező szótár*. Novotrade Kiadó, 1989.
- ³Russell, Stuart J.– Norvig, Peter: *Mesterséges Intelligencia - Modern megközelítésben*. Panem Könyvkiadó, Budapest, 2000. 26. fejezet, Online: http://project.mit.bme.hu/mi_almanach/books/aima/index [2015.06.08.]
- ⁴Csáji, Balázs Csanád: *A mesterséges intelligencia filozófiai problémái*, Szigorlati dolgozat, Eötvös Lóránd Tudományegyetem, filozófia szak. Budapest (2002) p. 4.
- ⁵Csáji, Balázs Csanád: im. p. 5.
- ⁶Platón: *Szofista*. ford.: Kövendi Dénes. Platón Összes Művei II., Gondolat kiadó, 1989.
- ⁷Descartes, René: *Értekezés a módszerről*, ford.: Szemere Samu, Boros Gábor. Ikon kiadó, 1993.
- ⁸Turing, Alan Matheson: *Computing Machinery and Intelligence*, (1950) *Mind*, 59, 433-460.
- ⁹Flanagan, Owen J.: *The Science of the Mind* (1992) 1-22. Idézi: Solum, Lawrence B.: *Legal Personhood for Artificial Intelligences*. North Carolina Law Review, Vol. 70, 1992. p. 1231.
- ¹⁰Hobbes, Thomas: *Leviatán* (1670) Idézi: Solum, Lawrence B.: im. p. 1234.
- ¹¹Russell, Stuart J.– Norvig, Peter: im.
- ¹²Searle, John R.: *Minds, Brains & Science*. 1984. pp. 28-41. Idézi: Solum, Lawrence B.: im. p. 1236.
- ¹³Moravec, Hans: *Mind Children: The Future of Robot and Human Intelligence*. Harvard University Press, 1988.
- ¹⁴Searle, John R. (1992). Idézi: Russell, Stuart J.– Norvig, Peter: *Artificial Intelligence: A Modern Approach* (2nd ed. Upper Saddle River, New Jersey: Prentice Hall, 2003.
- ¹⁵Mamoru, Osii: *Ghost in the Shell*. Production IG, Bandai Visual. 1995
- ¹⁶Solum, Lawrence B.: im. p. 1244.
- ¹⁷Kasparov és Deep Blue közötti győztes sakkmeccs leírása: http://en.wikipedia.org/wiki/Deep_Blue_versus_Kasparov,_1997,_Game_6 [2014.12.08.]
- ¹⁸Információk Deep Blue-ról: <http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/> [2014.12.08.]
- ¹⁹Allgrove, Benjamin D.: *Legal Personality for Artificial Intellecs: Pragmatic Solution or Science Fiction?* Master of Philosophy Thesis. University of Oxford. 2004. pp. 5-6.
- ²⁰Shelley, Mary: *Frankenstein, avagy a modern Prométheusz*. ford.: Göncz Árpád, Móra Könyvkiadó 1977. Elérhető: <http://mek.oszk.hu/02700/02735/02735.htm>
- ²¹Barthemess, Ulrike – Furbach, Ulrich: *Do We Need Asimov's Laws?* In: Lecture Notes in Informatics, Gesellschaft für Informatik, Bonn, 2014. p. 5.
- ²²Barthemess, Ulrike – Furbach, Ulrich: im.
- ²³Kaczynski, Theodore J.: *Az ipari társadalom és jövője*. Elérhető: http://unabomber.hu/az_ipari_tarsadalom_es_jovojje_utoszoval.pdf [2015.05.04.]
- ²⁴Dassanayake, Dion: *Bill Gates joins Stephen Hawking in warning Artificial Intelligence is a threat to mankind*. Express.co.uk, 2015. január 29. <http://www.express.co.uk/news/world/555092/Bill-Gates-Stephen-Hawking-Artificial-Intelligence-AI-threat-mankind> [2015.04.18.]
- ²⁵Kurzweil, Ray: *A szingularitás küszöbén*. Ad Astra Kiadó, 2014.
- ²⁶Russell, Stuart J.– Norvig, Peter: im. 26. fejezet
- ²⁷Russell, Stuart J.– Norvig, Peter: im. 26. fejezet
- ²⁸Russell, Stuart J.– Norvig, Peter: im. 25. fejezet
- ²⁹Asimov, Isaac: *Én, a robot*. Kossuth Könyvkiadó, 1966.
- ³⁰Asimov, Isaac: *Robotok és birodalom*. Móra Ferenc Ifjúsági Könyvkiadó, 1993.
- ³¹Hubbard, F. Patrick: „Do Androids Dream?”: *Personhood and Intelligent Artifacts*. Temple Law Review Vol. 83. 2010. pp. 85-86.
- ³²Asimov, Isaac: *Az első törvény*. Szukits Kiadó, 2001.
- ³³Sharkey, Noel: *Asimov törvényei már nem elegendőek*. 2010. szeptember 20. http://www.sg.hu/cikkek/77070/noel_sharkey_asimov_torvenyei_mar_nem_elegendoek [2014.12.03.]
- ³⁴Asaro, Peter M.: *Robots and Responsibility from a Legal Perspective*. HUMLab, Umea University, Sweden, 2007. Online: <http://www.peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf>
- ³⁵Sartor, Giovanni: *Cognitive Automata and the Law*. In: Artificial Intelligence and Law, Vol. 17, EUI Working Papers Law No. 2006/35.
- ³⁶Auer, Ádám – Papp, Tekla: *Rövid áttekintés a software fogalmáról a magyar szerzői jogban*. In: Jogelméleti szemle, 2013/2. pp. 10-17.
- ³⁷Petkó, Mihály: *A számítógépes programalkotások hatályos jogi védelme a nemzetközi jogalkotás tükrében*. In: Jogtudományi Közöny, 2002/1. pp. 53-54.
- ³⁸Pálos, György: *A szoftver és a szerzői jog*. In: *Ügyészek Lapja* 1996/2. szám
- ³⁹Tattay, Levente: *Az információs társadalom és az Európai Unió szerzői joga*. In: Tudományirányítás és a szellemi alkotások joga az EU-ban, MTA Tudomány és Kutatások Jogi és Igazgatási kérdéseinek Kutatócsoportja, Bp., 2003, p. 114.

-
- ⁴⁰ Lontai, Endre – Faludi, Gábor – Gyertyánfy, Péter – Vékás, Gusztáv: *Magyar polgári jog, Szerzői jog és iparjogvédelem*. Eötvös József Könyvkiadó, Budapest, 2012, p. 51.
- ⁴¹ Balogh, Zsolt György: *A jogi informatika válogatott fejezetei*. Janus Pannonius Tudományegyetem, ÁJK, Pécs, 1997, p. 45. és p. 54.
- ⁴² Dudás, Ágnes: *A szoftver szerzői jogi védelme I. rész*. In: Iparjogvédelmi és Szerzői Jogi Szemle, 2005/4. szám pp. 2-3.
- ⁴³ State of New Jersey 216th Legislature, Senate, No. 734. Online: http://www.njleg.state.nj.us/20142015/S1000/734_II.HTM [2015.04.15.]
- ⁴⁴ Az eredeti angol szöveg szerint: „Artificial intelligence means the use of computers and related equipment to enable a machine to duplicate or mimic the behavior of human beings.”
- ⁴⁵ Wein, Leon E.: *The Responsibility of Intelligent Artifacts: Toward an Automation Jurisprudence*. Harvard University Journal of Law and Technology, Volume 6, Fall Issue, 1992. pp. 105-111.
- ⁴⁶ Andrade, Francisco – Novais, Paulo – Machado, José – Neves, José: *Contracting Agents: Legal Personality and Representation*. Artificial Intelligence and Law, Vol. 15. 2007. p. 357.
- ⁴⁷ Russel, Stuart – Norvig, Peter: im., 2.1. és 27.1-2. fejezet.
- ⁴⁸ Kis, Sabrina: *Contracts and Electronic Agents: When Commercial Pragmatism and Legal Theories Diverge*, LLM Theses and Essays, Georgia Law 8-1. 2004. pp. 9-10.
- ⁴⁹ Kecskés, László: *Polgári jog. A személyek joga*. Dialog Campus, Budapest-Pécs, 2007.
- ⁵⁰ Idézi: Andrade, Francisco – Novais, Paulo – Machado, José – Neves, José: im. p. 359.
- ⁵¹ Idézi: Andrade, Francisco – Novais, Paulo – Machado, José – Neves, José: im. p. 359.
- ⁵² UNCITRAL Model Law on Electronic Commerce 13. § (2) bekezdés b)
- ⁵³ Dr. Petrik, Ferenc (szerk.): *Polgári jog. Kommentár a gyakorlat számára (negyedik kiadás)*. HVG Orac, Budapest, 2014. pp. 25-29.
- ⁵⁴ Sartor, Giovanni: im.
- ⁵⁵ Sztj. 42. § (1)-(2)
- ⁵⁶ Simon, Dávid: *A szoftverrel kapcsolatos egyes felelősségi kérdések*. In: Infokommunikáció és jog, 2005/3. HVG Orac Kiadó. p. 12.
- ⁵⁷ Dr. Petrik, Ferenc (szerk.): im. pp. 348-349.
- ⁵⁸ Dr. Petrik, Ferenc (szerk.): im., p. 350.
- ⁵⁹ Ptk. 6:174 (1) bekezdés
- ⁶⁰ Dr. Petrik, Ferenc (szerk.): im., p. 392.
- ⁶¹ A Ptk. 6:177. § (2) bekezdése alapján „a kötelezett mentesül a hibás teljesítés jogkövetkezményei alól, ha a hiba a) a jogosult által adott anyag alkalmatlanságára vagy hibájára; b) adat hiányosságára vagy hibájára; vagy c) utasítás célszerűtlenségére vagy szakszerűtlenségére vezethető vissza, és e körülményekre a kötelezett a jogosultat figyelmeztette.”
- ⁶² Dr. Petrik, Ferenc (szerk.): im., p. 392.
- ⁶³ Például az ún. Y2K szoftverhibákkal kapcsolatban lásd: Cégvezetés 19. szám: *A számítástechnikai cégek felelőssége*. Hírtőzsde Holding, 1999. október 1. Online: <http://cegvezetes.hu/1999/10/a-szamitastechnikai-cegek-felelossege/> [2015.05.16.]
- ⁶⁴ Ptk. 6:159. §
- ⁶⁵ Sztj. 59. § (1) bekezdés
- ⁶⁶ Ptk. 6:178. § (1) bekezdés
- ⁶⁷ Ptk. 6:152. §
- ⁶⁸ Dr. Petrik, Ferenc (szerk.): im., p. 937.
- ⁶⁹ Sartor, Giovanni: im.
- ⁷⁰ Sartor, Giovanni: im.
- ⁷¹ Mázi, András: *A veszélyes üzemi kárfelelősség magyarországi fejlődése a polgári korszakban*. Jogiforum.hu (2001.11.11.) Online: <http://www.jogiforum.hu/publikaciok/49.0.0> [2015.04.26.]
- ⁷² Ptk. 6:535. §
- ⁷³ Dr. Petrik, Ferenc (szerk.): im.. p. 981.
- ⁷⁴ Ptk. 6:536. §
- ⁷⁵ Dr. Petrik, Ferenc (szerk.): im.. p. 988.
- ⁷⁶ BH 1987.437
- ⁷⁷ EBH 2002.748
- ⁷⁸ Dr. Petrik, Ferenc (szerk.): im.. p. 986.
- ⁷⁹ Gabriel, Hallevy: *The Criminal Liability of Artificial Intelligence Entities* (2010. 02. 15.). Online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1564096
- ⁸⁰ Btk. 7. §
- ⁸¹ Földvári, József: *Magyar büntetőjog. Általános rész*. Osiris Kiadó, Budapest, 2006. pp. 98-99.

⁸² Ugo, Pagallo: *Killers, fridges and slaves: a legal journey in robotics*. In: *AI and Society*, Volume 26, Issue 4, pp. 347-354. Springer-Verlag, London, 2011.

⁸³ Btk. 3. §

⁸⁴ Lásd: A jogi személlyel szemben alkalmazható büntetőjogi intézkedésekről szóló 2001. évi CIV. törvény 1. §

⁸⁵ Földvári, József: im. p. 112.

⁸⁶ Lásd például Charlie Miller biztonságtechnikai mérnök alábbi előadását: Charlie Miller: *Car Hacking*, 10. Jubileumi Hacktivity, 2013. október 11. Online: <http://www.youtube.com/watch?v=kWiOVwP5GTE> [2014.01.07.]

⁸⁷ Btk. 423. § (1) c)

⁸⁸ „Emlékszel a kérdésre, ami arra készítette a Teremtőket, hogy megtámadjanak minket Tali’Zorah?: Ennek az egységnek lelke van?” A *Mass Effect* űrepszának harmadik részében a kollektív mesterséges intelligencia által irányított mechanikus geth fajt képviselő Legion teszi fel a kérdést, az azokat megalkotó quarian fajt képviselő Tali’Zorah-nak. A történet szerint a teremtőik ellen fellázadó, önálló tudatra ébredő gethek példájából kiindulva a galaktikus törvények szerint a mesterséges intelligencia-fejlesztés az egyik legsúlyosabb bűncselekménynek számít az egész galaxisban. Bioware, *Mass Effect 3* (Electronic Arts, 2012)