

**ARTIFICIAL INTELLIGENCE THROUGH
ASIMOV'S EYES
OR THE WORK OF A LIFETIME****A MESTERSÉGES INTELLIGENCIA
ASIMOV SZEMÉVEL
AVAGY EGY ÉLET MUNKÁJA**SZABÓ Lajos¹**Abstract**

Since 1956, countless devices and their programs have been declared to be artificial intelligence. The word „robot” has undergone a similar change of meaning since 1922. For 46 years, a science fiction writer wrote his romans and novels about the relationship between artificial intelligence and humans, at the end they was integrated into a coherent set of ideas at the end of his life. Today's scientists and developers are not concerned with the ethical aspects of AI, which Asimov's writings provide the main guidelines. In the 21st century, software is undergoing an amazing evolution, but the regulation of software is in its infancy, both ethically and legally.

Keywords

Asimov, AI, robots, ethical aspets of AI, ethical and legal regulation

Absztrakt

1956 óta számtalan eszközt és programjukat nyilvánították mesterséges intelligenciának. A "robot" szó is hasonló jelentésváltozáson ment keresztül 1922 óta. A sci-fi író 46 éven át írta novelláit és regényeit a mesterséges intelligencia és az ember kapcsolatáról, végül élete végén ezek egy összefüggő gondolathalmazba integrálódtak. A mai tudósok és fejlesztők nem foglalkoznak a mesterséges intelligencia etikai vonatkozásaival, amelyekhez Asimov írásai adják a fő iránymutatást. A XXI. században a szoftverek elképesztő fejlődésen mennek keresztül, a velük kapcsolatos szabályozás azonban gyermekcipőben jár, etikai és jogi értelemben is.

Kulcsszavak

Asimov, MI, robotok, az MI etikai vonatkozásai, etikai és jogi szabályozás

¹ szabo.lajos@uni-obuda.hu | ORCID: 0000-0001-9375-2188 | Lecturer Óbuda University / Institut of Safety Science and Cybersecurity | Chairman of the Board of Trustees, Foundation for Law Enforcement and Private Security Education and Research (REMOK) | kuratóriumi elnök, Alapítvány a Rendvédelmi és Magánbiztonsági Oktatásért és Kutatásért (REMOK)

A MESTERSÉGES INTELLIGENCIA ÉS A ROBOT KIFEJEZÉSEK EREDETE

Mielőtt Asimov életművének elemzésébe kezdek, szeretnék tisztázni néhány fontos kérdést. Meggyőződésem szerint fontos, hogy mindig pontosan tudjuk miről is beszélünk, ha egyes kifejezéseket használunk.

Mióta foglalkozunk a mesterséges intelligenciának nevezett dolgokkal?

„Sok kezdeti eredményt lehetne MI-nek nevezni, azonban egy teljes elképzelést az MI-ről 1950-ben Alan Turing fogalmazott meg a *Computing Machinery and Intelligence* c. cikkében. Itt vezette be a Turing-teszt, a gépi tanulás, a genetikus algoritmusok és a megerősítéses tanulás fogalmakat.”[1]

A második világháború után számos kutató kezdte el az öntanuló, döntési képességgel rendelkező programok kutatását, fejlesztését. Sokáig, sok néven nevezték ezeket mígnem 1956 nyarán, két hónapos munkatalálkozót nem szerveztek a témával foglalkozó tudósok. A soknevű témakör soknevű is maradt, amíg a Stanford Egyetem későbbi professzora John McCarthy a Dartmouth-i munkatalálkozón el nem nevezte, úgyhogy máig ezt a kifejezést használjuk.

Ahogy erről a Mesterséges Intelligencia Elektronikus Almanachban olvashatjuk:

„A munkatalálkozó talán legtartósabb eredménye az volt, hogy elfogadták a terület McCarthy által kreált új nevét, azaz a *mesterséges intelligenciát (artificial intelligence)*.”[2]

Ahogy ez előtt sokféleképpen nevezték az egyébként más-más problémakörökre tervezett eszközöket és a bennük működő programokat, azóta is annyiféle értelmezését találjuk a fejlesztők, kutatók, szakírók által alkalmazott mesterséges intelligencia fogalmaknak, leírásoknak.

Kiemelkedően fontos leszögezni, hogy a médiában elterjedt mesterséges intelligenciával kapcsolatos kifejezések, mint például „A mesterséges intelligencia már ezt is tudja.” A mesterséges intelligencia képes rá.” és hasonlók hibás képet alakítanak ki az olvasóban. Azt sugallják, mintha a mesterséges intelligencia egységes valami lenne, egy dologról beszélünk, ha róla beszélünk. Ennek a sugallatnak a következménye az, hogy a témában járatlan személyek is ebben a kontextusban használják. Pedig ez az értelmezés káros, megtévesztő és teljes mértékben tudománytalan, alaptalan.

Igen McCarthy óta van egy általános elnevezés, de az egymástól teljességgel elkülönült programokra és az azokkal irányított eszközökre vonatkozik. Teljesen mindegy, hogy egy szövegfeldolgozó- és vagy készítő programról beszélünk, amely szóban vagy írásban kommunikál velünk, vagy egy arcfelismerő szoftver dolgozik valamilyen hardverben, vagy egy sakkautomatával játszunk, az mind mesterséges intelligencia. Ugyancsak az a szóbeli vagy írásbeli kérésre képet, vagy mozgóképet generáló szoftver, vagy egy összetett rakéta-elhárító rendszer, négylábú állatok, vagy emberek különféle mozgásait elsajátító gépezeteket is egy részben öntanuló szoftver vezérli, és ezek mind mesterséges intelligenciák.

Azonban nyilvánvalóan semmi közük egymáshoz!

Amit az egyik tud, azt nem tudja a másik, még akkor sem, ha nyilvánvalóan vannak közös moduljaik is. A közös modulokban nem azonos szoftverek működnek, hiszen célra orientáltan hozták létre azokat más és más szakemberek a világ több kutató vagy fejlesztő laboratóriumaiban. Elegendő csak az információbevitelre gondolni. Nem mindegy, hogy mit és milyen spektrumban érzékel a beviteli modul és ezeket az érzéketeket milyen szoftver alakítja át észleletté. Hiszen ha az embert modellezzük, sokkal kisebb spektrumban érzékeli a

fényt, mint néhány állat, da ugyanígy vagyunk a hanggal a hővel és még sorolhatnánk. Ahogy a bolygónkon élő szervezetek mást és mást érzékelnek és azokra eltérően reagálnak, ugyanígy a sok-sok mesterséges intelligencia-változat is mást és mást érzékel és a programjának megfelelően reagál.

Azért nincs csak úgy általában mesterséges intelligencia és azért nem szabadna így beszélni róla, mert ahogy a különféle élő szervezetek más és más célra, más és más „szoftverrel” mint például a DNS, más és más bemeneti és kimeneti, információfeldolgozó, valamint végrehajtó eszközrendszerrel rendelkeznek, úgy minden mesterséges intelligencia programot használó rendszer is eltérő célokra készült.

E kutatások és fejlesztések mellett folyamatosan ment az automatizáció az iparban, ahol a monoton és nagy pontosságot kiváltó munkafeladatoknál az ember kiváltása volt a cél. Ez a folyamat napjainkban sem állt le, egyre jobb és jobb robotokkal rendelkezik az ipar és mára már az orvostudomány is. A XXI. században már számos otthoni automata berendezés végez munkát otthonainkban, és a számítógépeken, és egyéb eszközökön számos automata végzi el helyettünk a napi robotot, dolgozik helyettünk.

Ezzel meg is érkeztünk a következő problémához. Minden olyan automatát, mely valamilyen emberi munkát helyettesítő munkafolyamatot végez el, egy ideje robot-nak nevezük, pedig nem annak kellene, inkább automatának.

A robotok alapvetően nem olyanok kellene, mint azt manapság az ipari robotoknál és mindenütt másutt látjuk.

A robotok eredetileg ember formájú androidok voltak! Azokat és csak azokat nevezték robotnak, amelyek ember formájúak, autonóm mozgásra, kommunikációra képesek, érzékelni és az érzéketeket észleletté alakítani, vagyis információt létrehozni, és annak megfelelően adekvát cselekvésre képes homeosztátok voltak, melyek ráadásul öntanuló programmal is rendelkeztek.

Ugyanis a szó megalkotója, Karel Čapek, „Rossumovi univerzálni roboti” című színdarabjában egy olyan gépet ír le, ami ember formájú, és gondolkodni képes.

„A robot – a középkori jobbágyszolgáltatások közé tartozván – Magyarországon már korábban is ismert szó volt. Bár a szláv nyelvekben – így az oroszban – van általános jellegű „munka” értelmezése is a robotának, kényszermunka jelentésben is ismert. ... Bár Karel Čapek használta tényleg először a szót, nem ő találta ki. Egy angol nyelvű cseh weboldal közli Čapek visszaemlékezését milderre. A rövid történetből kiderül, hogy először „Labori”-nak akarta hívni a szerkezeteket, de a fivére javasolta neki a robot szót az ember alakú gépek megnevezésére.”[3]

Ennek a színdarabnak az 1922-es New-York-i bemutatója után kezdték el robotnak nevezni az ember formájú intelligens szerkezeteket, majd később mindenféle más gépi szerkezetre is alkalmazni kezdték, így jutottunk el a jelenkori értelmezéshez, amiről nyilván kevesen tudták eddig, hogy a szláv nyelvek munka-értelmű szavából ered.

Mivel a nyelv egy „élő” fejlődő kommunikációs rendszer, a benne használt szavak, kifejezések folyamatosan tartalmi, értelmezési szempontból átalakulni képesek. Az eredetileg a munkát jelentő szó, már a korai középkorban jelentett kényszermunkát, olyan munkavégzést, amelynek hasznából a munkát végző személy nem részesül. A földesúr számára végzett munka, amit a jobbágysági ellentételezés nélkül végeztek. Ez a kvázi rabszolgai tevékenység szintén a robot szóval gyökerezett meg több nyelvben, így magyarul is.

A Čapek-i értelmezésű robot, pedig nyilvánvalóan olyan rabszolga-automata, amelyik nem részesül gazdája érdekében végzett munkáért ellentételezésben.

A címben szereplő Isaac Asimov például tudta, hiszen le is írta az Alapítvány pereme című könyv 2. kötete, első fejezetében, ahol Quintesetz professzor beszélget Pelorat professzorral, és magyaráz neki a robotokról:

„Quintesetz összecücsörítette az ajkát, hátradőlt székében (mely enyhén megereszkedett a súlya alatt), és ujjhegyeit egymáshoz illesztette. Láthatóan gondot okozott neki, hogy hol is kezdje.

– Tudják, mi az a robot? – szólalt meg végül.

– Robot? – kérdezte vissza Pelorat. – Nem. Quintesetz Trevize-ra pillantott, aki lassan megrázta a fejét.

– És azt tudják-e, mi az a számítógép?

– Természetesen – felelte Trevize türelmetlenül.

– Nos, mozgó, számítógépes szerszám...

– Az, mozgó, számítógépes szerszám – fejezte be Trevize kissé ingerülten.

– Számptalan változata ismeretes, és nem tudok semmiféle általános érvényű kifejezést azon kívül, hogy mozgó, számítógépes szerszám.

– ... mely pontosan olyan, mint egy emberi lény: ez a robot – fejezte be S.

Q. nyugodtan a definíciót. – A robot abban különbözik a többitől, hogy ember formájú.

– Miért ember formájú? – kérdezte Pelorat őszinte elképedéssel.

– Ezt magam sem tudom. Szerszámnak határozottan alkalmatlan forma, ebben egyetértünk, én azonban csak a legendát ismételem. A “robot” régi szó, nem ered semmiféle felismerhető nyelvből, noha tudásaink szerint valami módon összefüggésben áll a „munká”-val.

– Egyetlen olyan szóról sem tudok – jegyezte meg Trevize kételkedve –, amely akár távolról is emlékeztetne a “robot”-ra, ugyanakkor bármiféle köze lehetne a “munká”-hoz.

– Bizonyára nem szerepel a galaktikus köznyelvben – hagyta rá Quintesetz –, mégis ezt állítják.

– Talán valami fordított etimológiáról lehet szó – vélte Pelorat. – Ezeket a tárgyakat használhatták munkavégzésre, s a szó lassanként átvette a “munka” jelentést...”[4]

Mint orosz születésű, Iszaak Judovics Ozimov, cirill betűkkel *Исаак Юдович Озимов*, aki 1920-ban született a Szovjetunióban az orosz föderáció Szmolenszk megyéjében található Petrovicsi faluban, nyilván beszélt valamennyire oroszul. Nyilvánvalóan a Čapek-i és a történelem során kialakult értelmezést is ismerte, hiszen a magyarázat tökéletesen erre mutat.

A MESTERSÉGES INTELLIGENCIA ASIMOV KOHERENS REGÉNYFOLYAMÁBAN

Asimovot, mint híres író, a legtöbbben szépíróként vagy tudományos-fantasztikus íróként és nem tudósként ismerik, vagyis elsődleges információ a nevéhez ez. Holott ismeretterjesztő tudományos műveinek száma is nagyon jelentős, valószínűleg meghaladja tudományos-fantasztikus műveinek a számát. Biokémikusként docensi fokozatot ért el, és

éveken keresztül tanított, tanszéket vezetett, mígnem úgy döntött, hogy szélesebb körben írásain keresztül folytatja a tanítást.

Saját magát olyan tudósként jellemezte, aki akkor is tudományos ismeretterjesztést folytat, amikor sci-fi műveket ír. Semmiképpen nem méltó arra, hogy bárki is legyintsen a műveiben leírtakra, vagy lekicsinyelje teljesítményét. Egy olyan szerző, aki tudományos fokozattal és tanári „előélettel” több mint 500 könyvet jegyez, és ezek jelentős része tudományos ismeretterjesztő munka, nem „csak egy sci-fi író”. Munkáit érdemes elolvasni és tudományos műveknél felhasználni. A jelen tanulmánynak nem célja irodalmi vagy részletes szövegszerű elemzéseket végezni, mindössze a legfontosabb mondanivaló kiemelése és ritkán, idézetekkel való igazolása fordul elő.

A jelen tanulmány Asimov azon írásait tekinti egy egységes gondolatfolyam részének, amit számos a témával foglalkozó irodalmár, kritikus és persze az írásait, mint a tanulmány szerzője is jól ismerő olvasója annak tekint. A novellák, kisregények és regények megjelenése nem időrendi, hiszen az Én a Robot-ban szereplő novellákat már az Alapítvány trilógia követte, majd az Űrvadász-sorozat, az Elijah Bayley-történetek, a Kavics az égen és a Csillagok akár a por, csak lazán kapcsolódik, de kapcsolódik, az Alapítvány előtt az Alapítvány pereme és az Alapítvány és a Föld már ezek után évtizedekkel, a kiadó felkérésére készült el. Asimov írói zsenialitása, hogy ha belekezdünk elejétől a végéig ebbe a hatalmas olvasnivalóba, egy egységes gondolatfolyam rabjai leszünk az elsőtől az utolsó mondatig.

Hát lássuk, miről is írt a világhírű író és tudós a mesterséges intelligenciával kapcsolatban.

Első írásaiban, melyben a „pozitronagy” kifejezést megalkotva, gondolkodó, éntudattal rendelkező, tanulni és dönteni képes automatákról ír Lehet, hogy a mai kiterjesztő robot értelmezést is neki köszönhetjük, hiszen már akkor is a robot kifejezést használja, amikor járművekbe képzelt pozitronagyat, így azokat gondolkodó, érzelmekkel és értelemmel rendelkező gépekké változtatja.

Asimovot 46 éven keresztül annyira foglalkoztatta a probléma, hogy csak az ember rendelkezik-e tudattal és önálló döntésre való alkalmassággal, hogy az első tudományos-fantasztikus írásaitól kezdve, majdnem az utoljára megjelent regényéig, fő mondanivalójául választotta. Minden bizonnyal örömmel csatlakozott volna a Cambridge-i Nyilatkozat aláíróihoz, ha megérhette volna.

Mi is az a Cambridge-i nyilatkozat? Csányi Vilmos: Etológia, ember, társadalom. című munkájában a következő módon magyarázza el röviden: „**Mindössze öt éve született meg az egyetértés a tudósok között arról, hogy az ember és az állat idegrendszere alapvetően egyezik, legfeljebb méretbeli és fajra jellemző különbségek léteznek. Vagyis a cambridge-i deklaráció kimondta: az állatoknak is van tudatuk, így minden állat gondolkodik, legfeljebb e gondolatok szerényebbek és más típusúak, mint az emberéi**”[5]

Ahogy a Cambridge-i nyilatkozat a tudatosságról, amit 2012. július 7-én adtak ki[6], a nem emberi élőlények – nem-emberi állatok kifejezést meggyőződésem szerint fordíthatjuk élőlényként - tudatállapotával kapcsolatban kijelenti azt, hogy;

„*A neokortex hiánya nem zárja ki, hogy a nem-emberi élőlények érzelmi állapotokat éljenek meg. Egybehangzó bizonyítékok arra utalnak, hogy a nem emberi élőlényeknél is megvannak neuroanatómiai, neurokémiai és neurofiziológiai alapjai a tudatos állapotnak, a szándékos viselkedés képességével együtt. Következésképpen a bizonyítékok súlya arra*

utal, hogy az ember nem egyedülálló a tudatosságot létrehozó neurológiai alapok birtoklásában. A nem emberi élőlények, beleértve az összes emlőst és madarat, és sok más élőlény, köztük a polipok is, szintén rendelkeznek ezekkel a neurológiai alapokkal"

Ehhez szeretném hozzátenni, hogy a tisztogatóhal (Labroides dimidiatus) úgy tűnik rendelkezik éntudattal, vagy valami csökevényével, mert felismeri magát a tükörben! [7] Úgy tűnik tehát, a Cambridge-i nyilatkozat a tudomány által szolgáltatott újabb felfedezések következtében akár már ennyi év eltelte után is kiegészíthető lenne.

Amennyire csak lehetett igyekeztem magyarítani a nyilatkozat végén található szöveget, a Neokortex kifejezést érintetlenül hagytam, csakúgy, mint a magyarul ideg- előtaggal általában fordított, neuro-előtagú szakkifejezéseket, mivel azok magyarításra eddig még a szakmai nyelvben sem teljesen elfogadott.

Hihetetlenül érdekesen kapcsolódik ehhez a XXI. századi nyilatkozathoz a XX. század talán legtermékenyebb írója Isaac Asimov, aki szinte teljes alkotói időszakában, (majdnem 50 éven keresztül!) vissza-vissza tért a tudattal kapcsolatos kérdésekhez. Élete végéig foglalkoztatta a mesterséges intelligencia, ezen belül a robotok fejlődésének elvi lehetősége, de a nem emberi és nem az ember által létrehozott mesterséges intelligencia is, mint arról később szövegek.

Az először 1964.-ben egy kötetben megjelent *Én a robot* [8] (I, ROBOT New York 1964.) majd később egy *Robottörténetek* [9] címen magyarul megjelent 1940-1976 között készített összes írása, ami a robotokkal kapcsolatos novellákat tartalmazza és egységes gondolatfolyam, pedig egyes elemei között hatalmas idő telt el.

Ezeknek a novelláknak a mondanivalója egy mesterséges értelem, a pozitronagyú, önálló éntudattal és gondolkodási, tanulási képességgel rendelkező, az embert kiszolgáló robotokról szól. Első ránézésre az ember formájú robotokról és más robotokról -járművek stb. - ír történeteket, melyek pozitronaggal, éntudattal és tanulási értelmi képességekkel rendelkeznek, a történetek a gazdáik és a közöttük folyó interakciókat érintik. Izgalmas és szórakoztató írások de a téma ismerőinek feltűnik, hogy mind a tudat, az emberi és mesterséges tudat/intelligencia problémája körül forog. A másik fontos téma az intuíció, de az nem tartozik a jelen témához, így mellőzöm.

Az általa leírt „Robotika három törvénye” nem csak logikai és döntési, hanem etikai és lélektani problémákat is felvet, melyek több évtizedes végig gondolása és leírása során, filozófiai mélységeikig jutnak könyveiben mind az emberek, mind a robotok.

Az *Úrvadász* sorozatban a nem emberi intelligencia és az emberi intelligencia találkozásának problematikáját feszegeti. Az általa elképzelt, különböző szinten értelmes létformák, melyek benépesítik galaxisunk égitesteit más-más kommunikációs csatornákat használva teremtenek kapcsolatot az emberekkel. Izgalmas gondolat kísérletek, filozófiai, erkölcsi problémafelvetésekkel.

A *Hajnal bolygó robotjai* [10] című regényében megismert Giskard aki a *Robotok és a Birodalom* című [11] regény végén R. Daneel Olivaw-val folytatott beszélgetést a három törvény kiterjesztéséről, a negyedik törvény bevezetéséről, szintén egy komoly etikai-filozófiai vita!

A robotpszichológia, mint kifejezés, megjelenik a legkorábbi novellagyűjteményben, az *Én a robot*-ban is, és a megalkotott kategóriától nem tágit későbbi írásaiban sem. A novellagyűjtemény külön fejezete az, amelynek a főszereplője, a robotokat gyártó monopólium, az Amerikai Robot neves robotpszichológusa, doktor Susan Calvin. A többi

főszereplő pedig természetesen mind robot, akik megjelenítenek olyan anomáliákat, melyek ma is jelen vannak, vagy bármikor megjelenhetnek a már létező alkalmazásokban.

Ezekben a novellákban majdnem mindent leírt, ami miatt ma félünk a mesterséges intelligenciától.

A szerző az egyszerű automaták világát jobbra elkerülve, a távoli jövőben feltalált „pozitronagy” segítségével, értelmes és egyre értelmesebb, szóban is kommunikáló robotokra alapozza az írásait. Egyértelműen a Capek-i robotok jelennek meg a történetekben. Nem tudhatjuk, hiszen külön nem írta le, de nyilvánvaló a problémakör ami ezek megírására készítette. Ez a mesterséges intelligencia és az emberi intelligencia, a gépi és emberi értelem, a gépi és emberi tudat volt az ami foglalkoztatta, abból a feltételezésből kiindulva, hogy egyszer megjelennek az ilyen képességekkel rendelkező robotok.

Az egyik novella, amit 1976-ban írt az Amerikai Egyesült Államok két évszázados fennállására, felkérésre, „A két évszázados ember” (angolul *The Bicentennial Man*)[12], kiemelkedő fontosságú. Eleinte arról szól, hogy egy ember formájú robot egy véletlen folytán művészi képességekkel rendelkezik és ezt a képességét gazdái örömmel veszik, támogatják művészi tevékenységét.

Megjegyzem a véletlen „programhiba” kétszer is előfordul ebben a novellasorozatban, mindkét esetben művészi képességek kialakulására tesz szert a robot. Ha ehhez hozzáteszem, hogy a különféle művészeti ágakban tevékenykedők képességei, az, hogy képesek érzelmeket, összetett gondolatokat, vagy akár egész történeteket színekben, vonalakban, foltokban, hangokban, táncban stb. megjeleníteni, az embereknél sem átlagos képesség. Van egy mondás, miszerint a zsenit és az örültet nagyon kevés választja el egymástól, a zsenihez nyugodtan a művészeket is hozzátehetjük. A két véglet ugyanis nyilvánvalóan olyan agyi, folyamatokban való különlegesség, mely az átlagtól eltér, a különféle „tehetségek” szintén az általánostól eltérőek, és mivel az agyunkban futó gondolkodási folyamatok következményei, tekinthetőek programhibának is.

E képességét felhasználva a robot műtárgyakat készít, melyek keresetté válnak a gyűjtők körében és a robotot valamint gazdáit híressé teszik. A művészeti alkotásokból származó bevételét első gazdája nem hajlandó átvenni, a pénzzel saját maga rendelkezik. A robotot első gazdája halálát követően, utódai összes nemzedéke hagyományosan családtagként és önálló entitásként kezeli, és ősükre emlékezve, a hagyományt megtartva önálló jövedelemmel rendelkezik. Egész hosszú, 200 éves életében folyamatosan tanul, és miközben annak a családnak az utódait szolgálja a család kihalásáig, melynek először a tulajdona lett. Végül tudományos kutatásokat végez, az emberi agy és a pozitronagy közti különbségeket és hasonlóságokat vizsgálja.

Először ruhákat kezd hordani, majd fémtestét az emberre megtévesztésig hasonlító mesterséges „emberi” kialakítású testrészekre cseréli. Ezek olyanok, amiket az emberek használnak protézisként és mesterséges szervekként. Miután minden testrészét ilyen félig bionikus testrészekre cseréltette ki, a pozitronagy kivételével, már csak egy vágya van. Bebizonyítani, hogy ezek után nincs különbség közte és az emberek között. Végül a robottörténetek főszereplője elismerteti magát embernek és ezt követően mint ember hal meg.

Az érzelmeket is megmozgató, lapos és jól felépített írás először mondja ki, annak a lehetőségét, hogy egyszer, valamikor, talán létrejöhet egy olyan összetett szoftver és a vele egybeépített hardver, mely szinte semmiben sem különbözik az emberektől. Igaz,

egyszeri és különleges kivétel még a regényben is, hogy ez a robot megkapja annak elismerését, hogy ember, és ezt is szinte a halála pillanatában.

Asimov a 46 éven keresztül írt és fokozatosan egységessé váló novella és regényfolyamban mindössze kétszer használja a mesterséges intelligencia kifejezést! A kifejezés, először a „Hajnal bolygó robotjai”[13] című könyve magyar kiadásának 58. oldalán a 3. bekezdés utolsó mondatában jelenik meg. Szövegszerűen:

„Baley igyekezett ráérezni a lényegre – a tendenciák, az általánosságok érdekelték –, és arra a megállapításra jutott, hogy az ember-robot kapcsolatban tapasztalható változások mind-mind a kölcsönös függőségi viszony kialakulása és megerősödése felé mutatnak. A robotok jogairól kötött megállapodás is ezt jelezte: fokozatosan megszűnik az, amit Daneel „felesleges megkülönböztetésnek” nevezett. Baley úgy vélte, hogy az auroraiak nem humanitárius meggondolásból viselkednek emberségesebben, hanem azért akarnak megfedkezni a robotok gépi mivoltáról, hogy ne zavarja őket a tudat: egyre inkább rá vannak utalva a mesterséges intelligenciára.”

A másik alkalom a regényfolyamban az „Az alapítvány előtt”[14] című regény 338. oldalán található, szövegszerűen:

„- Úgy látom, általános a számítógépesítés - jegyezte meg Dors. - Szerintem akár rá lehetne bízni az egész irányítást a komputerekre. Az efféle környezetben a legideálisabb a mesterséges intelligencia alkalmazása.”

Ez a két könyv eredetileg 1986-ban és 1988-ban jelent meg, sokkal később, mint a „mesterséges intelligencia kifejezés” amit, mint fentebb idéztem, 1956-ban a Drathmouthi munkatalálkozón alkottak meg és fogadtak el. Kifejezetten izgalmas, hogy bár biztosan ismerte, de megjelenésétől kezdve 30 éven keresztül került a szerző a kifejezés használatát, miközben nagyon sokat foglalkozott a témával.

Az Alapítvány Pereme[15] című könyvében, az általa elképzelt tudatos bolygó Gaia kapcsán, úgy írja le, hogy a bolygó minden élő és élettelen része valamilyen szintű tudatossággal rendelkezik a bolygómagtól a sztratoszféra utolsó bolygóhoz tartozó gázmolekulájáig. A bolygó minden része tudja, ismeri a helyét, a szerepét annak létezésében. Ezek a különféle szintű tudatformák érzékelik egymást és képesek az együttműködésre és a közös cselekvésre.

A témát a sorozat befejező kötetében is tovább gondolja, és az összes eddig leírt könyvét, novelláját e témában foglalja keretbe. Az Alapítvány és a Föld[16] végén egy olyan galaxist képzel el, mely egyéni tudatok közösségéből állva, közös tudatot alkot, Gaia mintájára a Galaxia a békés fejlődés és fennmaradás záloga.

Asimov majd fél évszázados munkája tehát messze túlmutat a jelen és közeljövő problémáin, és még sokkal jobban azokon, melyek megírásukkor voltak az akkori jelen és az akkori közeljövő problémái.

Pontosan látta az automaták és emberek együttélésének veszélyeit, az azokból adódó konfliktushelyzeteket, azoknak az emberi lélekre, tudatra való hatásait. Nem látta előre a szoftverek és a számítógépek közkézen forgását, a személyi számítógépek és tenyérben hordható változataik, melyeket első funkciójuk alapján máig telefonnak hívunk, holott már régen nem az a fő funkciója a legtöbbünk számára.

Azt azonban tökéletesen érezte, hogy eljön az a pont, ahol a gépek döntéseket hoznak majd az emberek helyett. Alapvetően nem a prediktív szövegbevitelre, vagy a helyesírás-ellenőrző, vagy a diktálás után szöveget lejegyezni képes programra gondolt – bár

ugye aki olvasta az Alapítvány trilógiát, tudja, hogy Arcadia Darell rendelkezik egy ilyen képességű leírószerkezettel – hanem az ennél sokkal mélyebben az ember helyett tevékenykedő szerkezetekre.

A Trantoron üzemelő automatikus vezérlésű légsiklók már átveszik az emberi irányítást, igaz bármikor vissza lehet venni az emberi ellenőrzést a jármű felett. Kicsit hasonlít a kezdetleges még csak félig „önvezető” járművekre, automata vészfékekre, sávtartó automatikákra, ESP, ABS és más megoldásokra, melyeket mi már a XXI. században mindennap használunk.

Mi azonban a XXI. században egyre pontosabban érzékeljük, mennyire oka lenne az Asimovi robotika három törvénye mindennapokban való alkalmazásának.

Emlékeztetőül:

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

A három törvény olyan filozófiai és etikai alapvetéseket rögzít, melyeket egy etikus gondolkodó és cselekvő embernek elvileg be kellene tartania saját életében, de be kell látnunk, hogy ez egy idealista elképzelés az emberek tekintetében. Még a „legszentebb” módon élő emberek is esetenként vétethetnek hibákat, és lássuk be, a három törvény a robotok alávetettségét fejezi ki az emberekkel kapcsolatban.

Tekinthetjük persze úgy is, hogy a robotika 3 törvénye azt célozza, hogy az egyes embereknek és az emberiségnek egészében semmi oka nincs tartania a mesterséges intelligenciától, hiszen alapvető programozása minden esetben kizárja, hogy veszélyt jelentsen az emberekre.

Ehhez a háromhoz kapcsolódik a már említett Nulladik törvény, amit először Daniel Robot Oliwaw mond ki, A Robotok és a Birodalom című könyv Párbaj című fejezet 63. sorszámú alfejezetében:

1. Egy robot nem árthat az emberiségnek, vagy nem veszélyeztetheti az emberiséget tétlensége által. [17]

A Nulladik törvény pedig olyan összetett és hatalmas számú összefüggést tartalmazó feladat, mely messze meghaladja az első három törvény könnyen és pontosan értelmezhető feltételeit. Olyan mennyiségű összetevőt kell figyelembe vennie, aminek csak az összeszámlálása is komoly időt venne igénybe. A rendelkezésünkre álló hardveres és szoftveres technológiával esetenként akár éveket, vagy évtizedeket is várni kellene egy olyan kérdés megválaszolására, mely a 0. törvény érvényre jutásával kapcsolatos.

Nyilvánvalóan a valóságtól kissé elrugaszkodott, de jó szándékú végletes idealizmus, és a regényfolyam további folytatásának vágya vezette erre Asimovot.

Azonban a gépeinktől és programjainktól, melyeket mi készítünk, magunknak, a mi hasznunkra, jó lenne, ha nem kellene félnünk.

ASIMOV ELVEI ÉS A VALÓSÁG

Igaz, még nem léteznek olyan mesterséges intelligenciák, melyek miatt alkalmazni kellene, sőt olyan robotok sem melyek ilyeneket használnának, de léteznek olyan robotok, és szoftverek, melyek az első törvény első fordulatában leírt tilalmat valamiképpen be kellene tartásuk.

A munkavédelmi előírások a mesterséges intelligenciával rendelkező robotok hiányában is az első törvény szellemében íródtak és működnek, a tervezés, létesítés, üzemeltetés, karbantartás során.

A takarító robotok akadály-észlelése és elkerülése szintén a sérülés okozását hivatott elkerülni.

A különféle házi automaták, mosó-, mosogató, szárítógépek ajtajai ugyanezért nem nyithatóak működés közben. Tehát a munkavédelmi és a hozzá kapcsolódó érintés- és tűzvédelmi szabványok megfelelnek az első törvénynek.

Nem sikerült azonban megfelelő védelmet létrehozni az ipari és háztartási kamerák tekintetében, mivel a hozzájuk kapcsolódó szoftverek készítői nem gondoltak arra, hogy kárt okozhat, ha nem fordítanak gondot arra, hogy sérülékeny ne legyen, vagy használóját figyelmeztesse.

Az a tény, hogy egyes számítástechnikai és IOT eszközök használhatnak olyan szoftvereket, melyek a használó, tulajdonos, vagy a szoftvert használó más személyek környezetében tartózkodó személyekről készítenek random felvételeket, melyeket eltárolnak, továbbítanak, megengedhetetlennek tűnik számomra.

Mindenképpen a robotika első törvényébe ütköznek azok a szoftverek, melyek bármilyen apró betűs részében deklarálják, hogy adatokat szereznek használóikról és azokat később felhasználják.

Tehát a személyes adatok védelme ugyanúgy beletartozik az első törvénybe, mint az élet, a testi épség, egészség védelme. Nyilvánvalóan az egészséget kiterjesztő módon értelmezve ide kell tartozzon a lelki egészség védelme is. Ez az, ami rögtön látszik, amikor egyes tartalmak megtekintéséhez, egyes eszközök használatához fondorlatos azonosítási megoldásokat igyekeznek kitalálni a fejlesztők, nehogy másnak a kezében használójára, vagy bárki másra veszélyt jelentsen.

Sajnos ki kell jelentenünk, kevés eredménnyel küzdünk a személyes adatok védelméért.

Számos olyan mesterséges intelligenciát fejlesztettünk ki, melyek akaratunk ellenére, megfigyelnek minket, árulkodnak rólunk, a kárunkra vannak.

Ugyanide, az első törvény hatálya alá kellene tartozniuk a különféle függőségeket okozó programoknak, vagy akár eszközöknek is, de erre végképp semmiféle korlátozásra való törekvés sem látható, sem a programok, eszközök készítői, sem a jogalkotók részéről!

Hány emberéletet mentett volna meg, ha a „telefon”- elnevezésű készülékek, melyeknek teljesítménye megegyezik esetenként egy számítógépével - programjai érzékelnék, hogy használójuk közlekedési helyzetekben veszélyben van és leállnának, vagy figyelmeztetnék a használót?

Tudunk olyan halálesetekről is, amikor a játékfüggő napokon keresztül ajzószerekkel pihenés nélkül játszott és belehalt függőségébe. Mindenki tudja, hogy a különféle betűkkel elnevezett korosztályok tagjainál sokszor elvonási tüneteket okoz megszokott eszközöktől akár egészen rövid ideig történő elválás. Sokszor hallottam, már többektől a

szakasztikus kifejezést, hogy „Ezeket már csak a gépek tartják életben!” és lássuk be van benne valami... Egyértelmű és nyilvánvaló pszichikai károkat okoz számukra, vagyis az első törvényt biztosan megsérti.

A virtuális valóságot létrehozó szerkezetek és programjaik is komoly defektusokat okozhatnak a függővé válók esetében.

Számos olyan programot terveztek és üzemeltetnek, mely az úgynevezett deep-fake[18] technológiát alkalmazva képes bárkiről képet[19], filmet[20], előállítani vagy akár a hangját[21] megtévesztésig hasonló módon utánozni. Ezekkel már komoly bűncselekmények is elkövethetők és nem kizárólag a személyes adatokkal visszaélés terén.

Igaz, ezek a készülékek sokkal kezdetlegesebbek az Asimov-i pozitronagyú robotknál, de ha nem is kényszeríti ilyesmire semmi a tervezőket, nem is fognak ilyen megoldásokra programokat készíteni!

Rengeteg olyan szerkezetet is előállított az emberiség, ami automatikusan kutat, érzékel, azonosít és megsemmisít eszközöket, tárgyakat vagy akár embereket., mint a rakétarendszerek, a Merkava Barak izraeli harcokcsirendszer, a drónhadviselés eszközei stb. Ezek lényegesen nagyobb hatásfokkal képesek azon műveletek önálló elvégzésére, mint a kizárólag emberi irányítással tevékenykednének.

A Gázai-övezetből indult támadás során a Hamasz egységei először a mesterséges intelligencia programokkal felszerelt automata határőr-tornyokat semmisítették meg. Ezek ugyanis nem csak a támadás, behatolás detektálását, hanem a támadók, behatolók megsemmisítését is elvégezték volna automata tűzvezető rendszerükkel.

Nézzük csak hol tartanak a szabályozás elképzeléseinél? Theodore Boone jogtanácsos és AI-szakértővel készült riport alapján még csak egyetlen aspektust vizsgálnak.

„EU az AI-ról szóló törvénytervezetében – amely még nem véglegesített vagy nem lépett hatályba, de úgy tűnik, hogy folyamatban van – azt a nézőpontot képviseli, hogy meg kell vizsgálnunk a különböző típusú AI-rendszerek által jelentett kockázatot, és az AI-rendszerek három kategóriáját kell létrehozunk...”

- *Az alacsony kockázatú mesterségesintelligencia-rendszerek minimális szabályozás és felügyelet alá esnének.*
- *A magas kockázatú mesterséges intelligenciával működő rendszerekre jelentős felügyeleti, átláthatósági és ellenőrzési követelmények vonatkoznának.*
- *A tiltott kategóriába sorolható mesterségesintelligencia-rendszerek közé tartozhatna például a nyilvános helyeken használt valós idejű arcfelismerő mesterségesintelligencia-technológia.”*[22]

Egyetlen szó sem esik az emberre veszélyes tevékenységekről, a háborús, rendőri és egyéb fegyverek alkalmazására képes AI által ellenőrzött rendszerekről. Ha vannak is ilyenek, azok nem nyilvános elvek és semmiképpen sem igazodnak az Asimovi „törvényekhez”. A tanulmány elkészülte óta az Európai Unió elkészítette és érvénybe léptette A mesterséges intelligenciára vonatkozó harmonizált szabályok, (A mesterséges intelligenciáról szóló jogszabály) megállapításáról és egyes uniós jogalkotási aktusok módosításáról szóló Európai Parlament és a tanács rendeletét, mely szóról szóra egyezik az ismertetett szakértői véleménnyel.[23]

Az alapelvek hézagosak, a szabályozás 3-5 éves időtartamokat is lehetővé tesz, mire alkalmazásra kerül, ami a technika és a programok fejlődési ütemét alapul véve azt jelenti, hogy a jogszabály, mire érvénybe lép a szabályozás biztosan elavulttá válik!

Az USA kormánya egy kicsivel sem jár előrébb e kérdésben. Számos irányelvet adtak ki elnöki rendelettel mint a Az amerikai vezető szerep megőrzése a mesterséges intelligencia területén[24], A megbízható mesterséges intelligencia használatának előmozdítása a szövetségi kormányban[25], vagy a Tervezet a mesterséges intelligencia jogairól szóló törvénytervezethez (Blueprint for an AI Bill of Rights)[26] ami egy jogszabály előkészítő anyag, amit 2020 óta igyekeznek használható formára alakítani. Meg kell jegyezni, eddig sikertelenül és a katonai, rendészeti alkalmazások kivételek, nem tartoznak e szabályozás tervezett köreibbe, pontosan ugyanúgy, ahogy az EU rendelet is kivételként kezeli ezeket. A tervezetben is vannak alapelvek, de azok meg sem közelítik az Asimovi 1. törvény szigorúságát.

MEGÁLLAPÍTÁSOK

Az etikai és jogi elvek és a napi gyakorlat bizony nem sokszor keresztezik egymás útját. Tudomásom szerint nincs etikai kódexe a szoftverfejlesztőknek, hardverkészítőknek, vagy például mechatronikai mérnököknek.

Mindazok a „felhasználási feltételek” melyeket az eszközökhöz, szoftverekhez kapunk – és amiket az emberek 99%-a olvasatlanul hagyva elfogad – majdnem teljesen a felelősség e- illetve áthárításáról szólnak, nem azoknak a védelméről, akik az eszközöket és szoftvereket használják.

Az Asimovi törvények, megítélésem szerint, etikus és alkalmas alapelvek lennének a szoftverek, hardverek és robotok, vagyis a mesterséges intelligenciát használó eszközök tervezésével foglalkozó emberek számára. Érdemes lenne elgondolkodni azon, hogy miért ódzkodik a szakma és a jogalkotók attól, hogy ilyen alapos alapelvek menték korlátozza mindazt, ami valószínűleg meghatározza századunk jelenét és jövőjét.

FELHASZNÁLT FORRÁSOK

- [1] Mesterséges Intelligencia Elektronikus Almanach TAMOP - 4.1.2-08/2/A/KMR-2009-0026 http://project.mit.bme.hu/mi_almanach/books/aima/ch01s03 1.3.1. A mesterséges intelligencia érlelődése (1943–1955) utolsó bekezdés
- [2] Mesterséges Intelligencia Elektronikus Almanach TAMOP - 4.1.2-08/2/A/KMR-2009-0026 http://project.mit.bme.hu/mi_almanach/books/aima/ch01s03 1.3.4. A mesterséges intelligencia megszületése 4. bekezdés utolsó előtti mondat.
- [3] A robot szó 90 éves 2011. január. 27. 13:45 https://hvg.hu/tudomany/20110127_90_eves_robot_szo_capek
- [4] Isaac Asimov: Az alapítvány pereme 2. 19-20. oldal, Kozmosz Könyvek HU ISSN 0324-5225, Foundation's Edge Doubleday and Co. Inc. Garden City New York 1982
- [5] Íme, az ember – Csányi Vilmos etológus az SZTE Mentor(h)áló program-sorozatában <https://u-szeged.hu/szتهirek/2017-aprilis/ime-ember-csanyi-vilmos>
- [6] A Cambridge-i nyilatkozat a tudatosságról. The Cambridge Declaration on Consciousness. <http://fcmconference.org/img/CambridgeDeclarationOnConsciousness.pdf>
- [7] Masanori Kohda, Takashi Hotta, Tomohiro Takeyama, Satoshi Awata, Hirokazu Tanaka, Jun-ya Asai, Alex L. Jordan: If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? <https://doi.org/10.1371/journal.pbio.3000021>

- [8] ÉN A ROBOT Kossuth Könyvkiadó 1966. fordította Vámosi Pál (I Robot, a signed book the new american library New York 1964
- [9] Isaac Asimov: Robottörténetek Móra Ferenc Könyvkiadó 1993 HU ISSN216-3244 ISBN 963 11 5 1. és 2. kötet, The complete robot Grafton an imprint of Harper Collins publishers 1983
- [10] Isaac Asimov: A hajnal bolygó robotjai Móra Ferenc Ifjúsági Könyvkiadó 1992 ISBN963 11 7011 X, The Robots of Dawn 1983 Doubleday and Co. Inc.Garden City New York
- [11] Isaac Asimov: A robotok és a birodalomMóra ferenc ifjúsági Könyvkiadó Rt. 1993. ISBN 963 11 7126 4 347. oldal, Robots and Empire Doubleday edition 1985 by Nightfall Inc.
- [12] Isaac Asimov: A hajnal bolygó robotjai Móra Ferenc Ifjúsági Könyvkiadó 1992 ISBN963 11 7011 X, The Robots of Dawn 1983 Doubleday and Co. Inc.Garden City New York
- [13] Isaac Asimov: Robottörténetek, Móra Ferenc Ifjúsági Könyvkiadó Rt. Budapest 1993, Második kötet 582-624. oldalak, The Complete Robot Grafton 1983
- [14] Isaac Asimov: Az Alapítvány előtt Móra Ferenc Ifjúsági Könyvkiadó 1991 ISBN963 11 6774 7, Prelude to Foundation 1988. by Nightfall Inc.
- [15] Isaac Asimov: Az Alapítvány pereme Kozmosz könyvek 1986 ISBN 963 211 680 1 Foundation's edge Doubleday and Comp. Inc. Garden City New York 1982
- [16] Isaac Asimov: Alapítvány és a Föld Móra Ferenc Könyvkiadó 1989 ISBN 963 11 6339 3Foundation and Earth Doubleday and Comp. Inc. 1986
- [17] Isaac Asimov: A robotok és a birodalomMóra ferenc ifjúsági Könyvkiadó Rt. 1993. ISBN 963 11 7126 4 347. oldal, Robots and Empire Doubleday edition 1985 by Nightfall Inc.
- [18] Az AI rátesz egy lapáttal a deepfake-re
<https://www.economx.hu/gazdasag/ai-summit-2023-mesterseges-intelligencia-aczel-petra-deepfake.777185.html>
- [19] Az AI sötét oldala: kislányokról generáltak meztelen képeket https://www.economx.hu/kulfold/mesterseges-intelligencia-gyermekpornografia-buncselekmény-spanyolorszag.777630.html?utm_source=index.hu&utm_medium=doz&utm_campaign=link
- [20] ‘Embrace it or risk obsolescence’: how will AI jobs affect Hollywood?
<https://www.theguardian.com/film/2023/aug/21/ai-jobs-hollywood-writers-actors-strike>
- [21] Élő énekesek hangját lopja el dalszerzéshez a mesterséges intelligencia
<https://www.economx.hu/gazdasag/zeneipar-mesterseges-intelligencia-deepfake-jogdj.775804.html>
- [22] Jönnek a robotok és velük a rideg valóság, hamarosan lefő a kávé <https://index.hu/gazdasag/2023/09/22/interju-mesterseges-intelligencia-theodore-boone-jog-munka-kitoresj-pont-sziget-egeszsegugy-radiologia/>
- [23] AZ EURÓPAI PARLAMENT ÉS A TANÁCS RENDELETE A MESTERSÉGES INTELLIGENCIÁRA VONATKOZÓ HARMONIZÁLT SZABÁLYOK (A MESTERSÉGES INTELLIGENCIÁRÓL SZÓLÓ JOGSZABÁLY) MEGÁLLAPÍTÁSÁRÓL ÉS EGYES UNIÓS JOGALKOTÁSI AKTUSOK MÓDOSÍTÁSÁRÓL https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/AUTRES_INSTITUTIONS/COMM/COM/2023/10-25/COM_COM20210206_HU.pdf

[24] Maintaining American Leadership in Artificial Intelligence. A Presidential Document by the Executive Office of the President on 02/14/2019 <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>

[25] Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, A Presidential Document by the Executive Office of the President on 12/08/2020 <https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government>

[26] Blueprint for an AI Bill of Rights <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>