

Négyesi Imre ezredes:

A MESTERSÉGES INTELLIGENCIA TÁRSADALMI ÉS ETIKAI KÉRDÉSEI¹

DOI: [10.35926/HSZ.2023.4.1](https://doi.org/10.35926/HSZ.2023.4.1)

ÖSSZEFOGLALÓ: *Miután a katonai tevékenységek során az emberi vezetés is kénytelen emberi életet kockáztatni, sőt feláldozni, a fő kérdés az, hogy mi lesz a különbség akkor, amikor ugyanezt a mesterséges intelligencia hajtja végre? Milyen társadalmi, etikai kérdéseket vethet fel a mesterséges intelligencia használata a mindennapi élet és kiemelten a katonai feladatok végrehajtása során?*

KULCSSZAVAK: *felelősség, magánélet, emberi jogok, a méltóság sérthetlensége, ellenőrzési mechanizmusok, hadi alkalmazás*

A SZERZŐRŐL:

Dr. habil Négyesi Imre ezredes (PhD), a Nemzeti Közszolgálati Egyetem Hadtudományi és Honvédtisztképző Kar egyetemi docense, az Informatika Tanszék vezetője (ORCID: 0000-0003-1144-1912)

TÁRSADALMI ÉS JOGI KÉRDÉSEK

A mesterséges intelligencia (MI) telepítésével kapcsolatos fő etikai dilemmák és erkölcsi kérdések áttekintésekor figyelembe kell venni a mesterséges intelligencia munkaerőpiaci hatásait, összpontosítva a gazdasági növekedésre és a termelékenységre valószínűsíthető hatásokra, a munkaerőre, a különböző demográfiai lehetőségekre gyakorolt lehetséges hatásokra, ideértve a digitális megosztottság és a mesterséges intelligencia munkahelyi alkalmazásának következményeit is. Vizsgálni kell, hogy miként lehetne megosztani a társadalmon belül a mesterséges intelligencia előnyeit, valamint azt is, hogy a technológia nagy internetes társaságokba történő koncentrálódása milyen kérdéseket vethet fel a politikai stabilitással kapcsolatosan. A vizsgálandó egyéb társadalmi kérdések közé tartoznak még a magánélet, az emberi jogok és a méltóság, az elfogultság és a demokrácia kérdései is.

A mesterséges intelligenciát használó robotok esetében, amelyek emberi társadalmi szerepet vállalnak (például ápolást), már figyelemmel kell kísérni az emberi pszichére gyakorolt hatásokat is. Ez számtalan kérdést vethet fel az ember–robot kapcsolatok vonatkozásában, illetve, hogy hogyan és milyen mértékben befolyásolhatják az ember–ember kapcsolatokat. Felvetődik ezáltal a személyiség kérdése is, és azt is vizsgálni kell, hogy kell-e rendelkezniük a mesterségesintelligencia-rendszereknek erkölcsi iránymutatással.

A mesterséges intelligenciának a pénzügyi rendszerre gyakorolt hatásai már a kezdetektől érzékelhetők. A pénzügyi szakértők kijelentették, hogy a mesterséges intelligencia felelős

¹ Részlet a szerző új kötetéből. Négyesi Imre: A mesterséges intelligencia katonai felhasználásának lehetőségei (Első kötet). Zrínyi Kiadó, Budapest, 2022.

a részvények magas kereskedési volumenéért. A pénzügyi kérdések mélyebb vizsgálata nélkül is nyilvánvaló, hogy noha a piacok alkalmasak az automatizálásra, vannak olyan kockázatok, amelyek között szerepel például, hogy a mesterséges intelligencia használata szándékos piaci manipulációhoz és összejátszáshoz vezethet.

Az új technológia használata kérdéseket vet fel mind a polgári, mind a büntetőjog szempontjából is, amelyek közül kiemelendő, hogy a meglévő jogi keretek vonatkoznak-e a mesterséges intelligenciák által hozott döntésekre. Felmerül tehát a felelősség kérdése, és bár valószínűtlennek tűnik, hogy a mesterséges intelligenciáknak megfelelő autonómiával és erkölcsi érzékkel kell rendelkezniük ahhoz, hogy maguk is felelőssé tehetőek legyenek, kérdéseket vetnek fel azzal kapcsolatban, hogy ki melyik bűncselekményért felelős. Az emberi szereplők a mesterséges intelligencia „mögé bújva” el tudják kerülni a felelősséget azzal, hogy azt állítják, nem tudják, hogy a mesterséges intelligencia képes lehet-e vagy megtenne-e egy ilyen dolgot. A felelősséggel kapcsolatos vitatott kérdések mellett vizsgálni kell azt is, hogy a mesterséges intelligencia elkövethet-e bűncselekményeket (például csempészést, mondjuk pilóta nélküli járművek használatával, zaklatást, kínzást, szexuális bűncselekményeket, lopást, csalást, bántalmazást stb.). Az önálló vezetésű autók valószínűleg újabb jogi kérdéseket vetnek fel a felelősséggel kapcsolatban, amelyek összetettebb eseteket eredményezhetnek (jelenleg a biztosítók általában elkerülik a pert, hogy meghatározzák, melyik járművezető a hibás, kivéve, ha egyértelműen kimutatható technikai hiba volt az autóban).

A mesterséges intelligenciák széles körű telepítése pozitív és negatív hatásokkal is járhat a környezetre, ezért ezt is vizsgálunk kell. A negatív hatások magukban foglalják a természeti erőforrások, például a ritkaföldfémek fokozott felhasználását, a szennyezés és a hulladék, valamint az energiafogyasztás növekedését. A negatív hatások mellett azonban meg kell említenünk azt is, hogy környezeti előnyökkel is járhat a használatuk, segíthetnek a hulladékgazdálkodásban és -megőrzésben.

2019-ben többször beszéltek a mesterséges intelligencia etikájáról, mint valaha. Több tucat szervezet készített etikai irányelveket, mielőtt még a médián keresztül felhívták volna ezek szükségességére a figyelmüket. Ma már nincs olyan mesterséges intelligenciával kapcsolatos konferencia, ahol a programozás kérdései mellett ne foglalkoznának az etikával is. Miután a kifejezetten a hadseregek számára fejlesztett mesterségesintelligencia-alkalmazások mellett a katonai szervezetek használnak egyéb MI-alkalmazásokat is, mindenképpen vizsgálunk kell a felhasználás etikai kérdéseit a katonai feladatok végrehajtásának vonatkozásában.

A mesterséges intelligencia fontosságát jól mutatja, hogy az Amerikai Egyesült Államokban a Fehér Ház tíz alapelvet fogalmazott meg az állami ügynökségek számára,² amelyeket be kell tartaniuk, amikor új, a magánszektorra vonatkozó mesterségesintelligencia-rendeleteket javasolnak. Az alapelveket a Fehér Ház Tudományos és Technológiai Politikai Hivatala (OSTP³)⁴ adta ki három fő cél figyelembevételével. A gyakorlatban az alapelvek kiadása után a szövetségi ügynökségeknek memorandumot kell benyújtaniuk az OSTP-nek, hogy elmagyarázzák azt, hogy a mesterséges intelligenciával kapcsolatos bármely javasolt rendelet hogyan felel meg az alapelveknek. Noha az iroda nem rendelkezik felhatalmazással a szabályok kiküszöbölésére, az eljárás továbbra is biztosíthatja a szükséges nyomást és koordinációt egy bizonyos szabvány fenntartása érdekében. R. David

² Justin Viktor: Az USA kiadta a mesterséges intelligencia tízparancsolatát. Raketa.hu, 2020. 01. 29. <https://raketa.hu/az-usa-kiadta-a-mesterseges-intelligencia-tizparancsolat> (Letöltés időpontja: 2022. 01. 18.)

³ Office of Science and Technology Policy.

⁴ <https://www.whitehouse.gov/ostp/> (Letöltés időpontja: 2022. 01. 18.)

Edelman, az MIT⁵ technológiai, gazdasági és nemzetbiztonsági projektigazgatója szerint megkísérelnek egy szabályozási előszűrést létrehozni, amelynek segítségével bizonyos szintű minőségellenőrzést vezethetnek be a mesterségesintelligencia-politikába.

A tíz alapelv a következőképpen hangzik:

- *A nyilvánosság bizalommal forduljon a mesterséges intelligencia felé.* A kormány minden eszközzel elősegíti a megbízható mesterségesintelligencia-alkalmazások létrehozását.
- *A nyilvánosság részvétele a létrehozásban.* A nyilvánosság számára lehetőséget kell biztosítani arra, hogy visszajelzést nyújtson a szabályalkotási folyamat minden szakaszában.
- *Tudományos integritás és információminőség.* A politikai döntéseknek minden esetben tudományosan megalapozottnak kell lenniük.
- *Kockázatértékelés és -kezelés.* Az ügynökségeknek el kell döntenüik, milyen szintű kockázatok a még elfogadható kockázatok.
- *Előnyök és költségek.* Az ügynökségeknek előzetesen mérlegelni kell valamennyi javasolt rendeletük várható társadalmi hatását.
- *Rugalmasság.* Bármelyik megközelítésnek alkalmazkodnia kell a mesterségesintelligencia-alkalmazások gyors változásaihoz és frissítéseikhez.
- *Igazságosság és megkülönböztetés-mentesség.* Az ügynökségeknek gondoskodniuk kell arról, hogy a mesterségesintelligencia-rendszerek ne diszkrimináljanak.
- *Közzététel és átláthatóság.* A nyilvánosság csak akkor bíz a mesterséges intelligenciában, ha tudja, mikor és hogyan kell használni.
- *Biztonság és védelem.* Az ügynökségeknek a mesterségesintelligencia-rendszerek által használt összes adatot biztonságos módon kell kezelniük.
- *Az ügynökségek közötti koordináció.* Az ügynökségek koordinációjával biztosítani kell, hogy következetesek és kiszámíthatóak legyenek a mesterséges intelligenciával kapcsolatos politikák.

ÁLTALÁNOS ETIKAI KÉRDÉSEK

A következő két idézet is bizonyítja, hogy miután a kifejezetten a hadseregek számára fejlesztettek mellett a katonai szervezetek használnak egyéb mesterségesintelligencia-alkalmazásokat is, ezért mindenképpen vizsgálnunk kell a felhasználás etikai kérdéseit a katonai feladatok végrehajtásának vonatkozásában. A fő veszély Yuval Noah Harari izraeli történész és író szerint: „A mesterséges intelligencia és gépi tanulás fejlődésével olyan tökéletes totalitárius rendszer építhető, amit még nem láttunk korábban.”⁶ Bár ez csak egy vélemény a sok közül, de a veszélyt valóban magában hordozza a mesterséges intelligencia. A történész a katonai alkalmazások tekintetében is eléggé szkeptikusan nyilatkozott: „Az önműködő fegyverek és gyilkos robotok technológiája veszélyesebb lehet akár a nukleáris fegyverkezésnél is, mert a nukleáris fegyverkezés esetében az irányítás kizárólagosan emberi döntések és együttműködés kérdése volt. A mesterséges intelligenciával olyan technológiát fejlesztünk, amelynek

⁵ Massachusetts Institute of Technology.

⁶ Kerner Zsolt – Kálmán Attila: Harari: A legkorruptabb politikusok a leghangosabb idegengyűlölők. 24.hu, 2019. 05. 09. <https://24.hu/belfold/2019/05/09/harari-orban-nacionalizmus-interju/> (Letöltés időpontja: 2022. 01. 18.)

irányítása nincs teljesen az emberiség kezében: ez a technológia képes lesz helyettesíteni az emberi elmét, és átvenni az irányítást az emberiségtől az élet számos területén.”⁷

Véleményem szerint tehát indokolt a katonai etikai kérdések felvetése. Ennek megfelelően először nézzük meg, milyen általános etikai kérdéseket vetett fel az eddigiek során a mesterséges intelligencia használata. Ezek között szerepelnek többek között:

- Hogyan védjük meg az emberek magánéletét?
- Hogyan dönthet a mesterséges intelligencia elfogulatlanul?
- Hogyan veszi figyelembe a mesterséges intelligencia a különböző emberek különböző értékrendjét?
- Ha a mesterséges intelligenciának annyi adata van szüksége, akkor hogyan biztosíthatók a személyiségi jogok?
- Hogyan kerül el a mesterséges intelligencia a nemek szerinti részrehajlást az ítéleteiben?
- Mit tekinthetünk a mesterséges intelligencia döntései közül igazságosnak?
- Hogyan bízhatunk a médiában az algoritmikusan létrehozott és elosztott dezinformációval szemben?

A már említett OSTP által kidolgozott tíz alapelvet minden ügynökségnek be kell tartania, amikor új, a magánszektorra vonatkozó mesterségesintelligencia-rendeletet javasol. Az alapelvek kiadásakor három fő célt vettek figyelembe:

- biztosítani a közvélemény részvételét;
- korlátozni a túlszabályozást;
- elősegíteni a megbízható, tisztességes, átlátható és biztonságos mesterséges intelligenciát.

A mesterséges intelligencia etikai kérdései vizsgálatának fontosságát a katonai alkalmazások tekintetében alátámasztja az Amerikai Egyesült Államok Védelmi Minisztériuma⁸ által „Summary of The 2018 Department of Defense Artificial Intelligence Strategy – Harnessing AI to Advance Our Security and Prosperity”⁹ címmel kiadott dokumentum, amelynek egy része külön foglalkozik a mesterséges intelligencia és az etika összefüggésével. Ebben a dokumentumban is vezető helyen szerepel a katonai etika kérdései között a konzultáció a hazai és a nemzetközi társadalom résztvevőivel (akadémia, magánipar, nemzetközi közösségek stb.). A dokumentum szerint az Amerikai Egyesült Államok Védelmi Minisztériuma arra is törekszik, hogy a mesterséges intelligencia felhasználásával csökkentsék a nem szándékos és a járulékos károkat a fokozott helyzet tudatosság és a fokozott döntéstámogatás révén. A technológia fejlesztésével és annak felhasználásával továbbra is megosztják céljaikat, etikai irányelveiket és biztonsági eljárásaikat, hogy ösztönözzék a nemzetgazdaság felelős fejlesztését és használatát más nemzetek számára is.

Az tehát látszik, hogy egyre több szinten, egyre több szervezetnél kezdtek vizsgálni a mesterséges intelligencia etikai, erkölcsi, társadalmi szempontjaival kapcsolatban. Először nézzük meg, mit mond a témában az Európai Bizottság független magas szintű mesterséges intelligencia

⁷ Rényi Pál Dániel: Harari a 444-nek: A klímaváltozással foglalkozzanak többet, ne a migrációval. 444.hu, 2019. 05. 09. <https://444.hu/2019/05/09/harari-a-444-nek-a-klimavaltozassal-foglalkozzanak-tobbet-ne-a-migracioval> (Letöltés időpontja: 2022. 01. 18.)

⁸ DoD: Department of Defense.

⁹ Summary of the 2018 National Defense Strategy of The United States of America. <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf> (Letöltés időpontja: 2022. 01. 18.)

szakértői csoportja,^{10, 11} az „Etikai útmutatás a megbízható mesterséges intelligencia számára” című dokumentumában.^{12, 13} A dokumentum első tervezetét 2018. december 18-án adták ki nyílt konzultációra, és több mint 500 közreműködőtől kaptak visszajelzést. A dokumentum az alapvető jogokon alapuló megközelítés alapján első fejezetében meghatározta az etikai alapelveket és az azokkal összefüggő értékeket, amelyeket tiszteletben kell tartani a mesterségesintelligencia-rendszerek fejlesztése, telepítése és használata során. A legfontosabb elvnek tartják az emberi autonómia tiszteletben tartását, a méltányosságot és a megmagyarázhatóságot. Fontosnak vélik, hogy felhívják a figyelmet a gyerekekre, a fogyatékosággal élőkre, a történelmileg hátrányos helyzetűekre és az olyan helyzetben levőkre, akiknél fennállhat a hátrányos megkülönböztetés lehetősége. A dokumentum rávilágít a mesterséges intelligencia használatának társadalmi kockázatára, valamint szükségesnek tartja a kockázat mértékével arányos intézkedések kiadását is. Külön pontokban sorolja fel a megvalósítás kulcsfontosságúnak ítélt alapelveit is, amelyek a mesterségesintelligencia-rendszerek teljes életciklusára vonatkoznak:

- teljes körű emberi felügyelet;
- műszaki megbízhatóság és felügyelet;
- a magánélet és a személyes adatok sérthetetlensége;
- átláthatóság;
- a környezeti és társadalmi jólét szem előtt tartása;
- elszámoltathatóság.

A legnagyobb technológiai cégek (mint a Google, a Facebook, a Microsoft, de még az Amazon is) gyakorlatilag rákényszerültek, hogy foglalkozzanak a mesterséges intelligencia etikai kérdéseivel. Ennek oka, hogy növekvő ellenállásba ütköztek a képzett munkaerők piacán, amióta egyre inkább hangot adnak a szoftverfejlesztők elégedetlenségüknek a cégek általuk vélt etikátlan együttműködési kapcsán. A tiltakozások egyik formája az volt, hogy néhány technikai dolgozó nyílt leveleket írt alá, amelyekben ellenzik többek között a Google Project Mavent¹⁴ és az amerikai katonasággal kötött szerződést. A Microsoft nem ért egyet az USA bevándorlási és vámügyi végrehajtásával, valamint az Amazon arcfelismerési technológiájának a bűnüldöző szerveknek történő eladásával. A cégek technikai alkalmazottai mellett már olyanok is csatlakoztak a kampányhoz, akik még nem köteleződtek el egyik nagy cég felé sem. Ők a nagy cégek fejadászainak küldött e-mailekben fejtik ki ellenérzésüket. A konkrét munkavállalók esetében a Tech Workers Coalition nevű munkajogi szervezet biztosít platformot azoknak a fejlesztőknek, akik anyacégeik működését etikátlannak tartják, és nem kívánnak részt venni az emberi jogi aggodalmakat kiváltó projektekből.

Az nyilvánvaló, hogy a szóbeli tiltakozás önmagában nem elég, mert sok szervezet mesterséges intelligenciával kapcsolatos etikai irányelve továbbra is homályosan fogalmaz és nehéz végrehajtani azokat. Kevés vállalat mutat be kézzelfogható változásokat a mesterségesintelligencia-termékek és -szolgáltatások értékelésének és jóváhagyásának módjában. A Google például létrehozott egy nominális mesterségesintelligencia-etikai tanácsot, amely azonban

¹⁰ Independent High-level Expert Group on Artificial Intelligence set up by the European Commission (AI HLEG). <https://ec.europa.eu/> (Letöltés időpontja: 2022. 01. 18.)

¹¹ Ethics Guidelines for Trustworthy AI. <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top> (Letöltés időpontja: 2022. 01. 18.)

¹² Uo.

¹³ Atomium European Institute. <https://www.eismd.eu/about/> (Letöltés időpontja: 2022. 02. 18.)

¹⁴ Az Egyesült Államok Védelmi Minisztériumának mesterségesintelligencia-projektje.

nem rendelkezik tényleges vétőjoggal a megkérdőjelezhető projektek felett, és csak pár tagja van, akiknek bevonása is vitát váltott ki. De 2019-ben eredményeket is sikerült elérni. Számos város – köztük San Francisco és Oakland (Kalifornia) és Somerville (Massachusetts) – betiltotta az arcfelismerés nyilvános használatát, és a javasolt szövetségi törvények hamarosan megtilthatják azt az amerikai állami intézményeknél is. A technológiai óriások (például a Microsoft, a Google és a Salesforce) alkalmazottai szintén egyre gyakrabban felhívták a figyelmet arra, hogy vállalataik mesterséges intelligenciát használnak a migránsok nyomon követésére és a drónok megfigyelésére.

Az érdekelt felek összehozására és a párbeszéd elősegítésére megalakult a „AI4People’s”, amely egy több felet képviselő fórum, amely összehozhatja az aktív innováció új alkalmazásai társadalmi hatásának kialakításában érdekelt összes szereplőt, beleértve az Európai Bizottságot, az Európai Parlamentet,¹⁵ a civil társadalmi szervezeteket, az ipart és a médiát. A hároméves ütemtervvel 2018 februárjában elindított „AI4People’s” célja egy közös nyilvános tér létrehozása az alapelvek, politikák és gyakorlatok meghatározására, amelyek alapján a „jó mesterségesintelligencia-társadalom” felépülhet. Ahhoz, hogy ez sikerrel járjon, meg kell állapodniuk abban, hogy miként lehetne legjobban ápolni az emberi méltóságot, elősegíteni az emberiség virágzását és gondoskodni egy jobb világról. Ez nemcsak a jogi elfogadhatóság, hanem az etikai preferencia kérdése is. A „The AI4People’s Ethical Framework for a Good AI Society” dokumentum¹⁶ a bevezetéssel együtt (1. szakasz) öt szakaszra épül. A 2. szakasz ismerteti a mesterséges intelligencia által kínált alapvető lehetőségeket az emberi méltóság és az emberiség virágzásának előmozdítására, valamint azok megfelelő kockázatait. A 3. szakasz rövid, magas szintű képet nyújt a szervezetek előnyeiről, ha etikai megközelítést alkalmaznak a mesterséges intelligenciák létrehozásakor. A 4. szakasz megfogalmazza a mesterséges intelligencia öt etikai elvét, a meglévő elemzésekre építve, amelyek alátámasztják a mesterséges intelligencia etikai elfogadását a társadalomban. Végül, az 5. szakasz 20 ajánlást kínál a „jó mesterségesintelligencia-társadalom Európában” kialakításának céljából.

A 4. szakaszban a szerzők szándékának megfelelően nem új alapelvek kerültek megfogalmazásra. Az „AI4People’s” célja pontosan az volt, hogy párbeszéd alakuljon ki az érdekeltek között – ezt segítené a kiadott dokumentum is. Az 5. szakasz ennek megfelelően a mesterségesintelligencia-társadalom ajánlásait ismerteti. Két részből áll: egy preambulumból és 20 akciópontból. A cselekvési pontok négy típusa létezik: felmérés, fejlesztés, ösztönzés és támogatás. Néhány ajánlást közvetlenül a nemzeti vagy európai döntéshozók tehetnek meg, adott esetben az érdekelt felekkel együttműködve. Mások számára a politikai döntéshozók hozzájárulhatnak a harmadik felek által tett vagy irányított erőfeszítésekhez. A szerzők határozott szándéka, hogy a mesterségesintelligencia-alapelvek beépüljenek a mesterséges intelligencia alapértelmezett gyakorlatába. Kiemelten fontosnak tartják, hogy a mesterséges intelligencia megmagyarázható legyen, mivel ez kritikus eszköz a nyilvánosság technológiába vetett bizalmának és megértésének felépítésében. Ennek megfelelően a 20 akciópont a következőképpen alakult:¹⁷

¹⁵ High-level expert group on artificial intelligence. <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence> (Letöltés időpontja: 2022. 01. 18.)

¹⁶ AI4People’s Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. <https://www.eismd.eu/wp-content/uploads/2019/03/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society.pdf> (Letöltés időpontja: 2022. 01. 18.)

¹⁷ Uo.

- Megvizsgálandó a meglévő intézmények, például a nemzeti polgári bíróságok azon képessége, hogy javítsák a mesterségesintelligencia-rendszerek által elkövetett hibákat vagy károkat.
- Mérjék fel, mely feladatok és döntéshozatali funkciók nem adhatók át a mesterségesintelligencia-rendszereknek.
- Megvizsgálandó, hogy a jelenlegi szabályozás kellően megalapozza-e az etikát ahhoz, hogy olyan jogi keretet biztosítson, amely képes lépést tartani a technológiai fejleményekkel.
- Hozzanak létre egy keretet a társadalmilag jelentős döntéseket meghozó mesterségesintelligencia-rendszerek megmagyarázhatóságának fokozására.
- Dolgozzanak ki megfelelő jogi eljárásokat az igazságszolgáltatási rendszer informatikai infrastruktúrájának fejlesztése, az algoritmikus döntések bírósági ellenőrzésének lehetővé tétele érdekében.
- Fejleszteni kell a mesterségesintelligencia-rendszerek ellenőrzési mechanizmusait a nem kívánt következmények elkerülésére (például tisztességtelen elfogultság azonosítása).
- Fejlesszenek ki jogorvoslati eljárást vagy mechanizmust a mesterséges intelligencia által okozott sérelmek vagy téves döntések orvoslására/kompenzálására.
- Fejleszteni kell a mesterségesintelligencia-termékek és -szolgáltatások megbízhatósági mutatóit.
- Új, a közjólét védelméért felelős EU felügyeleti ügynökség fejlesztése szükséges a mesterségesintelligencia-termékek, -szoftverek, -rendszerek vagy -szolgáltatások tudományos értékelése és felügyelete révén.
- A mesterséges intelligencia európai megfigyelőközpontjának kialakítása.
- Jogi eszközök és szerződéses sablonok kidolgozása az ember és gép közötti együttműködés megteremtésére.
- Ösztönözni kell pénzügyileg is az EU szintjén a mesterségesintelligencia-technológiák fejlesztését és használatát. Cél a társadalmi szempontból előnyösebb és sokkal inkább környezetbarát mesterséges intelligenciák létrehozása.
- A fenntartható, megnövekedett és koherens európai kutatási erőfeszítések ösztönzése pénzügyi szempontból is.
- A multidiszciplináris és ágazatközi együttműködés közötti vita ösztönzése a technológia, a társadalmi kérdések, a jogi tanulmányok és az etika kérdéseiben.
- A jogszabályok rendszeres felülvizsgálata annak tesztelésére, hogy milyen mértékben ösztönzik a társadalmilag pozitív innovációt.
- A mesterségesintelligencia-rendszerek empirikus tesztelése és fejlesztése.
- Ösztönözni kell a mesterséges intelligencia és annak alkalmazásai közvélemény általi felfogására és megértésére, valamint a strukturált nyilvános konzultációs mechanizmusok végrehajtására irányuló kutatások pénzügyi támogatását.
- A mesterséges intelligenciával kapcsolatos politikák és szabályok kidolgozása.
- Támogassa a mesterséges intelligenciával kapcsolatos szakmákra vonatkozó önszabályozó magatartási kódexek kidolgozását, magas etikai kötelezettségekkel.
- Támogatni kell a vállalati igazgatóságok azon képességét, hogy vállalják a felelősséget a vállalatok mesterségesintelligencia-technológiáinak etikai következményeiért.

Az etikai kérdések fontosságát bizonyítja Paolo Benanti ferences szerzetes is, aki régóta foglalkozik a mesterséges intelligenciával és annak etikájával. Már évek óta tevékenyen részt vesz a Vatikán mesterséges intelligenciára vonatkozó álláspontjának kialakításában,

sőt ő Ferenc pápa főtanácsadóinak egyike. Eredetileg mérnöknek tanult, és az ember és a technológiák találkozásának etikai vetületeiből írta doktori disszertációját. Annak érdekében, hogy összegyűjtse, mit tartanak etikusnak a nagy algoritmusfejlesztő cégek, személyesen találkozott többek között az IBM alelnökével, a Google-féle Alphabet tulajdonában álló DeepMind társalapítójával és a Facebook mesterségesintelligencia-etikával foglalkozó vezetőjével is. 2020. február 28-án Rómában a katolikus egyház képviselőjében a Pápai Életvédő Akadémia,¹⁸ a Microsoft, az IBM, az ENSZ Élelmezési és Mezőgazdasági Szervezete és az olasz innovációs minisztérium aláírta az azóta „Római felhívásként”¹⁹ ismertté vált dokumentumot, ami a mesterséges intelligencia etikus fejlesztésére szólít fel. Célja, hogy elősegítse a mesterséges intelligencia etikus megközelítését. Az eredeti ötlet az volt, hogy előmozdítsák a nemzetközi szervezetek, kormányok, intézmények és a magánszektor közötti közös felelősségérzetet egy olyan jövő megteremtése érdekében, amelyben a digitális innováció és a technológiai fejlődés központi szerepet tölt be az emberiség számára. Rámutatva egy új algoritmusra, az aláírók vállalták, hogy olyan mesterséges intelligencia fejlesztését kérik, amely az emberiség egészét, minden embert szolgál, amely tiszteletben tartja az ember méltóságát, hogy minden egyén részesülhessen a technológia vívmányaiból, és ennek nem az egyetlen célja a nagyobb profit vagy az emberek fokozatos helyettesítése a munkahelyen. A „Római felhívás” a mesterséges intelligencia etikájára három hatásterületet és hat alapelvet tartalmaz.

A három hatásterület:

- *Etika*: minden emberi lény szabadnak és egyenlőnek születik méltóságában és jogaiban.
- *Oktatás*: a világ átalakítása a mesterséges intelligencia innovációján keresztül azt jelenti, hogy jövőt építünk a fiatalabb generációk számára.
- *Jogok*: a mesterséges intelligencia fejlődésének az emberiség és a bolygó szolgálatában tükröznie kell az embereket – különösen a gyengéket és a hátrányos helyzetűeket – és a természeti környezetet védő előírásokat és elveket.

A hat alapelv:

- *Átláthatóság*: a mesterségesintelligencia-rendszereknek mindenki számára érthetőnek kell lenniük.
- *Befogadás*: ezek a rendszerek nem diszkriminálhatnak senkit, mert minden embernek egyenlő a méltósága.
- *Elszámoltathatóság*: mindig kell lennie valakinek, aki felelősséget vállal azért, amit egy gép csinál.
- *Pártatlanság*: a mesterségesintelligencia-rendszerek nem lehetnek részlehajlóak és nem hozhatnak létre torzításokat.
- *Megbízhatóság*: a mesterséges intelligenciának megbízhatónak kell lennie.
- *Biztonság és adatvédelem*: ezeknek a rendszereknek biztonságosnak kell lenniük, és tiszteletben kell tartaniuk a felhasználók magánéletét.

A „Római felhívást” teljes egészében a Vatikán szponzorálta, és május közepén Abu-Dzabiban a muszlim és a zsidó vallás képviselői is aláírták – reményeik szerint vallásokon átívelő lesz. Egy interjúban Paolo Benanti arra a kérdésre, hogy mit gondol, az algoritmusfejlesztés

¹⁸ Pontifical Academy for Life.

¹⁹ Call for AI Ethics.

melyik területén van a legnagyobb szükség az etikai iránymutatásokra, azt válaszolta, hogy valószínűleg azoknál az algoritmusoknál, amelyek ítéletet hoznak az emberek felett. Ebbe a sorba tipikusan beleillenek a katonai célú mesterségesintelligencia-fejlesztések, ezért egy külön fejezetben tekintjük át a katonai etikai kérdéseket.

KATONAI ETIKAI KÉRDÉSEK

A legfontosabb kérdés az, hogy a mesterséges intelligencia sikere az emberi faj végét jelentheti-e? Számtalan sci-fi-író fantáziáját foglalkoztatta régen is a kérdés, hogy a robotok vagy a robot-ember kiborgok rendezhetnek-e ámokfutást? Gyakorlatilag bármely technológiában benne rejlik a lehetőség, hogy kárt okozzon rossz kezekben, de a mesterséges intelligencia és a robotika esetében ezek a rossz kezek akár magához a technológiához is tartozhatnak. A fő veszélyt sokan abban látják, hogy a robotok nélkülözhetetlenek lesznek az ember számára és öntanuló képességükkel olyan magas szintű intelligenciára tehetnek szert, amely már meghaladja a természetes emberi intelligenciát.

A robotok viszonyát a társadalmi együttélés szabályaihoz, az erkölchöz és az átalakításhoz először egy DARPA-kutatáson keresztül nézzük meg.²⁰ A mesterséges intelligencia fejlődése révén a virtuális és robotsegédek egyre jobban képesek lesznek a bonyolult feladatok elvégzésére. Ahhoz, hogy ezeket az intelligens gépeket biztonságosnak, megbízhatónak, együttműködőnek, esetleg még etikusnak is lehessen tekinteni, a robotoknak képesnek kell lenniük arra, hogy gyorsan felmérjék az adott helyzetet és alkalmazzák az emberi-társadalmi normákat. Az emberi együttélés szabályainak megtanítása a robotoknak természetesen egy új kihívást jelent.

Ennek a kihívásnak a kezelésére az Egyesült Államok Védelmi Minisztériumának kutató-sokért felelős részlege, a DARPA²¹ által finanszírozott kutatók végrehajtottak egy projektet, amelynek célja egy elméleti és formális keret biztosítása volt a normák és a normatív hálózatok számára. Kísérletileg tanulmányozták, hogyan jelenítik meg és aktiválják a normákat az emberi elmék, és megvizsgálták, hogyan lehet megtanítani ezeket a normákat az algoritmusok segítségével a robotoknak. A csoport létrehozott egy kognitív számítási modellt az emberi normákról, és kifejlesztett egy gépi tanulási algoritmust, amely lehetővé teszi a gépek számára, hogy normákat tanuljanak ismeretlen helyzetekben az emberi adatokra támaszkodva.

Ghanadan Reza, a DARPA programmenedzsere azt mondta, hogy „[j]elenleg nem tudjuk, hogyan lehet az értelmes normafeldolgozást beilleszteni a hatékony számítási architektúrákba”, és hozzátette, hogy a társadalmi és etikai normáknak számos olyan tulajdonságuk van, amelyek egyedülálló kihívást jelentenek. Az emberek a normákat rendkívül gyorsan aktiválják, vagyis nagyon gyorsan észlelik a normák megsértését! A normák értelmezése azonban az emberi szubjektumtól is függ, és ez rendkívül megnehezíti az emberi normák gépi tanulásának vizsgálatát. „Ahhoz, hogy egy robot etikus is legyen, képesnek kell lennie számos olyan norma megtanulására, reprezentálására, aktiválására és alkalmazására, amelyeket egy adott társadalomban az emberek elvárnak” – mondta Ghanadan Reza. Ez a feladat sokkal bonyolultabb, mint a mesterségesintelligencia-rendszerek szabályainak megtanítása egyszerűbb feladatokra, például képek címkézésére, spam észlelésére vagy az emberek adóbevallásának elkészítésére. Azáltal, hogy keretet biztosít ilyen komplex algoritmusok

²⁰ <https://www.darpa.mil/news-events> (Letöltés időpontja: 2022. 01. 18.)

²¹ Defense Advanced Research Projects Agency.

kifejlesztéséhez és teszteléséhez, az új kutatás hamarabb elhozhatja azt a napot, amikor a gépek az emberi viselkedés legjobb módszereit követik. Összefoglalva Ghanadan Reza szavait, kimondhatjuk, hogy ha ugyanolyan közel kerülünk egymáshoz a jövőbeni robotokkal, ahogyan azt ma elképzeljük, akkor ezeknek az asszisztenseknek ugyanazokat a normákat kell betartaniuk, mint nekünk.

Az erkölcsileg kompetens robotok kutatása természetesen nem csak a DARPA-nál folyik. A tudósok érdeklődését egyre inkább felkelti, hogy milyen módon tehető egy robot „érző emberi lényé”. Néhány kérdés, amelyre a kutatások folyamán összpontosítottak:

- Hogyan lehetne beépíteni egy robotba az emberi normarendszerek néhány alapvető tulajdonságát?
- Milyen különbségeket találunk az ember és a robot között a döntések során betartott erkölcsi normák tekintetében?
- Milyen alapvető erkölcsi szókinccset kell megtanítani a robotnak, amely lehetővé tenné a mesterséges ágensek számára az erkölcsi események megismerését és az azokkal kapcsolatos kommunikációt?
- Hogyan használhatja fel a robot a megszerzett információkat saját cselekedeteinek irányításához?
- Hogyan ítélik meg az emberek az erkölcsi döntéseket meghozó robotokat?

Az erkölcsi normák tulajdonságaival, valamint az ember és a robot döntéseinek vizsgálatával kapcsolatosan számtalan kutatási eredményt közzétettek. Sok tanulmány foglalkozott azzal a kérdéssel is, hogy az emberek miként viszonyulnak az erkölcsi döntéseket meghozó robotokhoz. A tapasztalatokat összegezve arra a következtetésre jutottak, hogy az emberek kétharmadának nincs gondja azzal, hogy a robotokat vagy a mesterséges intelligenciát erkölcsi döntéshozókként kezelik. Kimutatták azt is, hogy amikor az emberek erkölcsileg értékeli a mesterséges ágenseket, ugyanazokat a normákat alkalmazzák, mint az emberek esetében. Érdekes eredmény továbbá, hogy az emberek néha kevésbé hibáztatják az embereket (vagyis önmagukat) olyan esetekben, amikor a döntés erkölcsileg indokolható. Igaz ez akkor is, ha a robot ugyanerre a döntésre jutott a tényleges probléma megoldásánál. Például az emberek kevésbé hibáztatják az embereket abban az erkölcsi döntést igénylő helyzetben, amikor az emberek nem hajlandók feláldozni egy egyént négy ember megmentése érdekében. Ez egyértelműen abból adódik, hogy az emberek megértik és a saját gondolataikban szimulálják, milyen nehéz lenne egy ilyen áldozatot hozó döntést meghozatala. Ez a fajta „gondolati szimuláció” és megértés nem érhető el egy robot műveleteinek megfontolásakor és kiértékelésekor.

A jövőbeni robotcsapatokban a sikeres interakciók támogatásához szükségük van arra, hogy a robot „emberi bizalmat” érezzen csapattársa iránt. Ez a projekt abból indult ki, hogy két vagy több partner közötti együttes fellépés megköveteli a viselkedés összehangolását, amely azonban többek között az erkölcsi normákra is támaszkodik. Azok a partnerek, akik betartják ezeket a normákat, megbízhatóknak minősülnek. Az emberek, akik a robotokkal közösen feladatot látnak el, ennek megfelelően arra számítanak, hogy „robotpartnereik” ismerik és követik a megfelelő normákat. Ha a robotok megsértik ezeket, magyarázatot kell adniuk és igazolniuk kell viselkedésüket az emberek számára érthető módon. (Ez a projekt tehát összekapcsolódik a robot viselkedésének magyarázatával kapcsolatos projektekkel.)

A mesterséges intelligenciát használó fegyverrendszerekre különös figyelmet kell fordítanunk, tekintettel arra, hogy a katonai tevékenységek súlyos károkat okozhatnak a társadalom

számára. A kérdés az, hogy a mesterséges intelligencia alkalmazása súlyosítja vagy enyhíti-e az okozott károkat? A napvilágra került esettanulmányok feltárják a mesterséges intelligencia hadviselésben való használatát és a mesterségesintelligencia-alkalmazások fegyverként való felhasználásának lehetőségét. Idetartoznak a halálos autonóm fegyverek, a dróntechnológiák, a robotgyilkosság és a mobil-robot improvizált robbanóeszközök.

A mesterséges intelligencia szélesebb körű katonai felhasználása során felmerülő kulcsfontosságú etikai kérdések közé tartoznak az emberi ítélet bevonásával kapcsolatos kérdések, mert ez sértheti a nemzetközi humanitárius jogot. Ha a mesterséges intelligencia fokozott használata csökkentené a háborús küszöböt, az befolyásolná a globális stabilitást.

Bár a katonai technológiában legalább a második világháború óta részben autonóm és intelligens rendszereket alkalmaznak, a gépi tanulás és a mesterséges intelligencia fejlődése fordulópontot jelent az automatizálás hadiüzemben való alkalmazásában. A mesterséges intelligencia már elég fejlett és kifinomult ahhoz, hogy olyan területeken alkalmazható legyen, mint például a műholdas képek elemzése és a kibervédelem, de az alkalmazások valódi felhasználási körét még nem sikerült teljes mértékben megvalósítani. Egy közelmúltbeli jelentés arra a következtetésre jutott, hogy a mesterségesintelligencia-technológia képes a hadviselést ugyanolyan vagy akár még nagyobb mértékben átalakítani, mint a nukleáris fegyverek, repülőgépek, számítógépek és a biotechnológia megjelenése.²²

A következőkben nézzük meg, hogy a mesterséges intelligencia milyen hatással lesz a haderőre. Széles körű bevezetésével valószínűleg folytatódni fog a folyamat, hogy a haderő egyre inkább átadja a hatalmat a technikának, amely a mesterséges intelligencia által inspirált fegyverkezési versenyhez vezethet. Az orosz katonai ipari bizottság már jóváhagyott egy tervet, amely szerint az orosz haderő 30%-a teljesen távvezérelt és autonóm robotplatformokból áll majd 2030-ra. Más országok is valószínűleg hasonló célokat tűznek ki. Míg az Egyesült Államok Védelmi Minisztériuma korlátozásokat vezetett be a halálos erőt gyakorló autonóm és félig autonóm rendszerek használatáról, más országok és nem állami szereplők nem gyakorolnak ilyen önmegtartóztatást.

A döntések és tervek háttérében elsősorban pénzügyi megfontolások állnak. A szokásos katonai repülőgépek több mint 100 millió dollárba kerülhetnek, ellenben egy kiváló minőségű kisméretű pilóta nélküli légi jármű jelenleg körülbelül 1000 dollárba kerül, ami azt jelenti, hogy egy csúscategóriás repülőgép áráért a katonaság egymillió drónt szerezhet be. Noha a jelenlegi kereskedelmi drónok korlátozott hatótávolsággal rendelkeznek, a jövőben hatótávolságuk a ballisztikus rakétákhoz hasonló méretű lehet, így elavulttá válnak a meglévő platformok. Az olcsó, nagy teljesítményű, halálos és autonóm robotok széles körű elérhetősége már jelentős etikai kérdéseket vethet fel. A szélesebb körben elterjedt és könnyebben hozzáférhető drónok lehetővé teszik a célzott gyilkosságot.

Mivel a kereskedelmi robotok és autonóm járműtechnológiák elterjedtek, néhány csoport ezt kihasználja, hogy korszerűbb, továbbfejlesztett robbanókészülékeket készítsen. Jelenleg a robbanóanyagok gyors célba juttatása sok mérföld távolságba csak a nagyobb nemzetállamokra jellemző technológiai képesség. Ha azonban a drónon keresztül hosszú távú csomagszállítás valósággá válik, a távolról történő pontosan leadott robbanóanyagok költsége a millió dolláros nagyságrendről ezres vagy akár százas nagyságrendre is csök-

²² Gregory C. Allen – Taniel Chan: Artificial Intelligence and National Security. Harvard: Belfer Center Study, 2017.

kenhet. Hasonlóképpen, az önjáró autók gyakoribbá és pusztítóbbá tehetik az öngyilkos autóbombákat, mivel már nem igényelnek öngyilkos vezetőt.

Wael B. Hallaq²³ már 2017-ben kiemelte azokat a kulcsfontosságú területeket, ahol a gépi tanulás valószínűleg befolyásolja a hadviselést. Leírt egy példát, amely szerint a parancsnok alkalmazhatja egy intelligens virtuális asszisztenst egy csatatéri környezetben, amely automatikusan beolvassa a műholdas képeket az egyes járműtípusok felismerése érdekében, segítve a veszélyek előzetes azonosítását. Megjósolhatná az ellenség szándékát is, és összehasonlíthatná a helyzeti adatokat a korábbi háborús harcok tárolt adataival, amely feladatot eddig egy tapasztalt és képzett törzs hajtott végre, biztosítva a parancsnok számára a felhalmozott tudást a vezetés végrehajtásához.

A mesterséges intelligencia hadviselésben történő alkalmazása számos jogi és etikai kérdést vet fel. Az egyik aggodalom tehát az, hogy az emberi megítélést kizáró automatizált fegyverrendszerek megsérthetik a nemzetközi humanitárius jogot; veszélyeztethetik az élethez való alapvető jogunkat és az emberi méltóság elvét. A nemzetközi humanitárius jog tágabb értelemben a nemzetközi jognak az a területe, amelyik a fegyveres konfliktusokkal és az azok során tanúsított magatartások szabályozásával foglalkozik (*ius in bello*). A mesterséges intelligencia ugyan csökkentheti a háborús küszöböt, de hatással lehet a globális stabilitásra, ezért jelentős hatást gyakorolhat a nemzetközi humanitárius jog szabályozási céljának elérésére, amely szerint cél a fegyveres ellenségeskedések káros hatásainak csökkentése, elsődlegesen a nem harcolók, a sebesültek, a polgári személyek, a hadifoglyok védelme, illetve az okozott károknak az éppen még szükséges szintre szorítása.

A nemzetközi humanitárius jog tehát előírja, hogy minden támadásnak különbséget kell tennie a harcosok és a nem harcosok között, arányosnak kell lennie, és nem szabad polgárokat vagy polgári tárgyakat céloznia. Ezenkívül egyetlen támadás sem súlyosbíthatja feleslegesen a harcosok szenvedését. Előfordulhat, hogy a mesterséges intelligencia nem képes ezeket az elveket teljesíteni emberi döntés bevonása nélkül. Egy olyan autonóm katonai robottípus, amely a halálos erő segítségével önállóan kereshet és „bevonhat” célokat, esetleg nem felel meg a nemzetközi humanitárius jog által meghatározott normáknak, mivel nem képes megkülönböztetni civileket a harcosoktól, ezáltal nem tudná megítélni, hogy a támadás erőarányos-e az okozott polgári károk tekintetében.

FELHASZNÁLT IRODALOM

- AI4People’s Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. <https://www.eismd.eu/wp-content/uploads/2019/03/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society.pdf>
- Allen, Gregory C. – Chan, Taniel: *Artificial Intelligence and National Security*. Harvard: Belfer Center Study, 2017.
- Atomium European Institute. <https://www.eismd.eu/about/>
- Ethics Guidelines for Trustworthy AI. <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>
- High-level expert group on artificial intelligence. <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

²³ Wael B. Hallaq az Avalon Alapítvány professzora (Columbia Egyetem).

- Independent High-level Expert Group on Artificial Intelligence set up by the European Commission (AI HLEG). <https://ec.europa.eu/>
- Justin Viktor: *Az USA kiadta a mesterséges intelligencia tizparancsolatát*. Raketa.hu, 2020. 01. 29. <https://raketa.hu/az-usa-kiadta-a-mesterseges-intelligencia-tizparancsolat>
- Kerner Zsolt – Kálmán Attila: *Harari: A legkorruptabb politikusok a leghangosabb idegengyűlölők*. 24.hu, 2019. 05. 09. <https://24.hu/belfold/2019/05/09/harari-orban-nacionalizmus-interju/>
- Négyesi Imre: *A mesterséges intelligencia katonai felhasználásának lehetőségei (Első kötet)*. Zrínyi Kiadó, Budapest, 2022.
- Office of Science and Technology Policy. <https://www.whitehouse.gov/ostp/>
- Rényi Pál Dániel: *Harari a 444-nek: A klímaváltozással foglalkozzanak többet, ne a migrációval*. 444.hu, 2019. 05. 09. <https://444.hu/2019/05/09/harari-a-444-nek-a-klimavaltozassal-foglalkozzanak-tobbet-ne-a-migracioval>
- Summary of the 2018 National Defense Strategy of The United States of America.
- <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>
- <https://www.darpa.mil/news-events>

A mesterséges intelligencia katonai felhasználásának lehetőségei

A technológia, valamint a mesterséges intelligencia jelentette filozófia kiemelkedően fontos a 21. században, korunk hadseregeiben is, hiszen a mesterséges intelligenciára alapozott katonai vezetési rendszerek vagy akár a fegyverirányítási rendszerek mind hatékonyságunkban, mind sikerességükben már ma is bizonyítják e technológia létjogosultságát, és ez prognosztizálható már a közeli jövőre is. A kötet alapos összefoglalót ad a mesterséges intelligencia elméleti hátteréről, majd megvizsgálja számos ország mellett a NATO stratégiai elképzelését erről a forradalman új területről.

Szerző: Négyesi Imre
Megjelenés éve: 2022
kartonált
232 oldal

3150 Ft



A könyv a Zrínyi Kiadó webshopjában (shop.hmzrinyi.hu) vagy a kiadó könyv- és térképboltjában (1024 Budapest, Filler utca 14.) vásárolható meg.