



---

NORTHWESTERN UNIVERSITY SCHOOL OF LAW  
KÖZJOGI ÉS JOGELMÉLETI SOROZAT - NO. 10-12

---

**A mesterséges  
intelligencia  
felgyorsítása**

*Northwestern University Law Review, hamarosan  
megjelenik*

**John O. McGinnis**  
**Northwestern University School of Law**

## GYORSULÓ AI

*John O. McGinnis*\*

A közelmúltban a mesterséges intelligencia ("AI") a média érdeklődésének középpontjába került. Tavaly májusban például a *New York Times* egy cikket szentelt annak a kilátásnak, hogy mikorra a mesterséges intelligencia eléri, majd felülmúlja az emberi intelligenciát.<sup>1</sup> A cikk az ilyen esemény és az "erős mesterséges intelligencia" veszélyeiről elmélkedett.<sup>2</sup> Júliusban aztán a *Times* a számítógépek által vezérelt hadviselésről értekezett. Különböző szakértők aggodalmuknak adtak hangot a számítógépek növekvő hatalmával kapcsolatban, különösen akkor, amikor azok új fegyverek alapjává válnak, mint például a ragadozó drónok, amelyeket az Egyesült Államok most terroristák megölésére használ.<sup>3</sup>

Ezek a cikkek összefoglalják a mesterséges intelligenciával kapcsolatos kettős félelmet, amely szabályozásra készítheti ezt a területet - az emberek által ellenőrizhetetlenné váló gépektől való egzisztenciális félelmet és a gépek pusztító erejétől való politikai aggodalmat egy forradalmasított csatatéren. Mindkét félelem túlzó. Az egzisztenciális félelem azon a téves elképzelésen alapul, hogy az erős mesterséges intelligencia szükségszerűen az emberi rosszindulatot tükrözi. A katonai félelem azon a téves elképzelésen alapul, hogy a számítógépek által vezérelt fegyverek szükségszerűen rontják, nem pedig mérséklék az emberi rosszindulatot.

Mindenesetre, tekintettel a

---

\*Stanford Clinton Sr. jogászprofesszor, Northwestern Egyetem. Köszönöm Mark Movsesian, Michael Rappaport és Ardith Spence hasznos észrevételeiket. Nagyon hálás vagyok Michael Abramowicznak is a témáról a Law and Accelerating Technology című közös szemináriumunkon folytatott megbeszéléseinkért.

<sup>1</sup> Lásd John Markoff, *Superbrain*, NYTIMES.COM

*The Coming*

,május23,2009,http://www.nytimes.com/2009/05/24/weekinreview/24markoff.html ([link](#)).

<sup>2</sup> *Id.*

<sup>3</sup> *Lásd* John Markoff, *Scientists Worry Machines May Outsmart Man*, NYTIMES.COM, július 25, ttp://www.nytimes.com/2009/07/26/science/26robot.html (2009, [link](#));

NYTimes.com, Predator Drones and

UnmannedAerialVehicles,http://topics.nytimes.com/topic/reference/timestopics/subjects/u/unmanned\_aerial\_vehicles/index.html (utolsó látogatás 201029, márciusában) ([link](#)).

a számítógépes teljesítmény növekedésének központi szerepét a katonai technológiában, lehetetlen lenne szabályozni a mesterséges intelligenciával kapcsolatos kutatást anélkül, hogy a világ legrosszabb nemzeteit felhatalmaznánk.

A mesterséges intelligencia betiltása vagy szigorú szabályozása helyett az Egyesült Államoknak támogatnia kellene a civil kutatást egy olyan mesterséges intelligenciával kapcsolatban, amely nem veszélyezteti az emberiséget.

emberek - egy úgynevezett "barátságos mesterséges intelligencia".<sup>4</sup>

Először is, ez a támogatás a legjobb módja annak, hogy biztosítsuk, hogy

hogy a számítógépek nem bizonyulnak egzisztenciális fenyegetésnek. A világ legfejlettebb technológiájú nemzetének kutatói számára ösztönzőket biztosítana az emberbarát mesterséges intelligencia kutatására és fejlesztésére.

Másodszor, ez a támogatás a számítási eredmények által a kollektív döntéshozatalban várhatóan jelentkező pozitív hatások miatt indokolt. A technológia felgyorsulása szükségessé teszi, hogy a kormányzat gyorsabban reagáljon a bomlasztó innovációk potenciálisan hatalmas hatásaira. Például annak a korszaknak a hajnalán, amikor az energiaigényes gépek feltalálása elkezdhetette felmelegíteni a Földet, kevesen ismerték fel az ilyen gépek által okozott magasabb hőmérsékletből eredő kockázatot.<sup>5</sup> Mégis, amint azt alább leírom, a technológia jelenlegi fejlődése miatt az energiaigényes gépek térhódítása lassúnak tűnik. Feltételezve, hogy az ember alkotta

<sup>4</sup> A "barátságos mesterséges intelligencia" meghatározását lásd: SINGULARITY INSTITUTE FOR ARTIFICIAL INTELLIGENCE, CREATING FRIENDLY AI § 1, <http://www.singinst.org/upload/CFAI.html> (utolsó látogatás: 2010. március 29.), amely a barátságos mesterséges intelligencia céljait úgy foglalja össze, hogy a mesterséges intelligencia az "önkéntelen fájdalom, halál, kényszer és ostobaság" megszüntetésére törekszik. ([link](#)). Javasolnék egy még gyengébb meghatározást, amely egyszerűen azt

biztosítaná, hogy az AI ne okozzon kárt az embereknek, és ne korlátozza szabadságukat sem rosszindulatból, sem ostobaságból.

<sup>5</sup> Igaz, hogy egy svájci kémikus az 1890-es években - még jóval azután is, hogy a iparosodás - hogy az emberi tevékenység globális felmelegedést okozhat. *Lásd Bradford C. Mank, Standing and Global Warming: Is Injury To All Injury to None?* , ENVTL35. L. (2005)1.,12 Mindazonáltal még az 1930-as években is különként tartották azt a tudóst, aki megismételte ezt az aggodalmat. *Lásd Maxine Burkett, Just Solutions to Climate Change: A Climate Justice Proposal for a Domestic Clean Development Mechanism*, BUFFALO56 L. REV. n.169,1733 (2008) ([link](#)).

a légkör felmelegedése zajlik,<sup>6</sup> valószínűleg csak az első a gyorsuló technológiai változások által előidézett lehetséges katasztrófák közül, amelyek megelőzhetők vagy legalábbis enyhíthetők korábbi objektív elemzéssel és figyelmeztetéssel. Nem kevésbé fontos azonban annak felismerése, hogy más technológiai fejlesztések a társadalom számára előnyök kaszkádját hozhatják létre - olyan előnyöket, amelyeket a kockázat téves megítélése hátráltathat, vagy akár ki is zárhat. Ennek eredményeképpen az információk gyors összegyűjtése és elemzése fontosabb, mint valaha a demokratikus döntéshozatalban, mivel a szabályozási döntések tétje még soha nem volt ilyen nagy.

Tekintettel arra, hogy a mesterséges intelligencia jelentős potenciállal rendelkezik abban, hogy segítse a társadalmat abban, hogy megfelelő politikát alakítson ki az összes többi, átalakító képességgel rendelkező, gyorsuló ütemben fejlődő technológiával - például a nanotechnológiával és a biotechnológiával - kapcsolatban, a technológiai változások szempontjából a legfontosabb politika magával a mesterséges intelligenciával kapcsolatos politika. Az erős mesterséges intelligencia segítene elemezni az adatokat minden szempontból.

a világ exponenciális ütemben növekvő adataiból.<sup>7</sup> AI akkor segíthet, hogy a szakpolitikák és a következmények közötti kapcsolatok, amelyeket az emberek máskülönben figyelmen kívül hagynának, tűzjelzőként működve az új technológiákból eredő veszélyekkel szemben, amelyek hatásláncolatát nehéz lehet felmérni, még akkor is, ha történelmi szempontból meglehetősen közel állnak.

Az ilyen elemzés nemcsak a katasztrófa elkerülése érdekében hasznos, hanem a felgyorsuló technológiából származó előnyök bőséges tárházának kihasználása érdekében is. A jövőbeli következmények jobb elemzése segíthet a kormánynak

abban, hogy a legjobb politikát alakítsa ki az ilyen jótékony technológiák támogatására, beleértve a megfelelő díjak és támogatás biztosítását is.

---

<sup>6</sup> Nem szükséges biztosnak lenni abban, hogy az ember okozta légköri változások bekövetkeznek ahhoz, hogy felismerjük, hogy hasznos lenne egy ilyen esemény kockázatát felmérni. Amint azt később javaslom, a mesterséges intelligencia egyik előnye, hogy segít az ilyen kockázatok értékelésében.

<sup>7</sup> Lásd ECONOMIST.COM, *The Data Deluge*, 2010. február 25., [http://www.economist.com/opinion/displaystory.cfm?story\\_id=15579717](http://www.economist.com/opinion/displaystory.cfm?story_id=15579717) (a modern világban rendelkezésre álló adatok gyors növekedésének leírása) ([link](#)).

fejlesztés. Talán még ennél is fontosabb, hogy a technológiai fejlődés hatásainak jobb elemzése csökkenti a technológiai változások által gyakran kiváltott félelmeket. Minél jobban elemezzük a jelenlegi technológia jövőbeli következményeit, annál kevésbé valószínű, hogy az ilyen félelmek elfojtják a hasznos innovációkat, mielőtt azok a prométheuszi fejlődést hozhatnák.

Ebben a rövid esszében először azt írom le, hogy miért van jelentős esélye annak, hogy az erős mesterséges intelligencia valósággá válik, majd felvázolom azt a két veszélyt, amelyet egyesek a mesterséges intelligenciának tulajdonítanak.

Megmutatom, hogy a mesterséges intelligenciáról való lemondás vagy annak hatékony szabályozása a versengő szuverén államok világában nem tud hatékonyan reagálni ezekre a fenyegetésekre, mivel a szuverén államok katonai előnyre tehetnek szert a mesterséges intelligenciából, és mivel még az államokon belül is nagyon nehéz lenne megakadályozni, hogy az egyének mesterséges intelligenciával kapcsolatos kutatásokat végezzenek. Sőt, azt javaslom, hogy a mesterséges intelligencia által vezérelt robotok a csatatéren valójában kevesebb pusztításhoz vezethetnek, és civilizáló erővé válhatnak a háborúban, valamint a civilizáció segítségére lehetnek a terrorizmus elleni küzdelemben.

Végezetül pedig felhozom, hogy miért lehet barátságos mesterséges intelligenciát fejleszteni, amely inkább segíti, mint károsítja az emberiséget, és így kiküszöböli az egzisztenciális fenyegetést.

Befejezésül bemutatom, hogy a tilalmi vagy szigorú szabályozási rendszerrel ellentétben a mesterséges intelligenciát támogató, a barátságosság elveit követő kormányzati politika a legjobb megközelítés a mesterséges intelligenciával kapcsolatban. Ha kialakul a barátságos mesterséges intelligencia, az segíthet



megelőzni az erős mesterséges intelligencia kevésbé barátságos változatainak kialakulását, valamint megkülönböztetheti a valódi veszélyeket a felgyorsuló technológia más formáiban rejlő számos potenciális előnytől.

## I. AZ AI ELJÖVETELE

A mesterséges intelligencia gondolata, amely elég erős ahhoz, hogy beavatkozzon az emberi ügyekbe, a *2001: Űrodüsszeia* HAL-jától kezdve a *Wall-E* robotjaiig a tudományos fantasztikum tárgya volt.<sup>8</sup> Az emberi intelligenciával vetekedő, sőt azt felülmúló számítógépek gondolata elsősorban inkább tűnhet spekulatív fantáziának, mint olyan témának, amelynek a jogi elemzés és politika napirendjén kiemelt helyet kellene kapnia.<sup>9</sup> A Holdra való utazás azonban egykor maga is a tudományos fantasztikum egyik alapműve volt a tizenkilencedik és huszadik században.

évszázado k.<sup>10</sup> Mégis, egyetlen kormányzati program miatt, az ember felfedezése a Földön

hold immár több mint negyvenéves történelmi esemény. És a holdraszállással ellentétben a mesterséges intelligencia fejlődése közvetlen hatással van a társadalmi kormányzásra.

Az erős mesterséges intelligencia az absztrakt gondolkodás és problémamegoldás általános emberi képességével rendelkező gépek létrehozása. Általánosan elfogadott tény, hogy ha az ilyen gépek lehetségesek, akkor hamarosan felülmúlják az emberi kognitív képességeket, mivel ugyanazok a folyamatok, amelyek létrehozták őket, gyorsan fejleszthetik azokat. Maguk a gépek is segíthetik ezt a folyamatot azzal, hogy az embernél nagyobb mértékben képesek megosztani egymás között az információkat.<sup>11</sup>

Az erős mesterséges intelligencia sikere három feltételezés igazságától függ. Az első premissza a funkcionalizmus. A funkcionalizmus azon a tételen alapszik, hogy a megismerés

---

<sup>8</sup> *Lásd:* 2001: A SPACE ODYSSEY (Metro-Goldwyn-Mayer 1968); WALL-E (Pixar Animation Studios 2008).

<sup>9</sup> A mesterséges intelligencia figyelembevétele a jogi tudományban nem kapott nagy teret. Egy érdekes cikk azt elemzi, hogy a mesterséges intelligencia képes-e a vagyonkezelő szerepét betölteni. *Lásd* Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, N70.C. L. REV. 1231 (1992).

<sup>10</sup> *Lásd pl.* JULES VERNE, A FÖLDTŐL A HOLDRA (1865).

<sup>11</sup> *Lásd* Irving John Good, *Speculations Concerning the First Ultraintelligent Machine*, ADVANCES IN COMPUTERS 31-3631, (1965) ([link](#)).

elkülönül attól a rendszertől, amelyben a megismerés megvalósul.<sup>12</sup> Így az absztrakt gondolkodás egyaránt megvalósulhat egy biológiai rendszerben, például az agyban, vagy egy elektronikus rendszerben, például egy számítógépben. E hipotézis szerint a szimbólumok rendszere, ha egy fizikai folyamat által megfelelően aktualizálódik, "képes intelligens cselekvésre".<sup>13</sup>

John Searle filozófus a legjelentősebb azon tudósok között, akik megkérdőjelezi azt az elképzelést, hogy egy absztrakt szimbólumokat manipuláló gép az emberi elmével egyenértékűvé válhat. Searle egy kínai szoba analógiáját hozza fel.<sup>14</sup> Ha valakit betesznek egy szobába, és kínaiul tesznek fel neki kérdéseket, akkor írásos utasításokat kaphat arra vonatkozóan, hogyan manipulálja a kínai írásjeleket, hogy kínaiul válaszoljon a kérdésekre.<sup>15</sup> Ő maga azonban semmit sem ért a kínai nyelvből, és ennek következtében a szimbólumok manipulálása az emberi megértés gyenge szimulációja.<sup>16</sup> Searle analógiájával szemben az egyik erős ellenvetés az, hogy az egész rendszer - az írott utasítások és az emberi manipulátor - érti a kínai nyelvet.<sup>17</sup> Searle tehát igazságtalanul antropomorfizálja a megértés tárgyát. Amint azt alább kifejtem, a számítógépes intelligenciáról való gondolkodás legnagyobb rendszerszintű hibája, hogy összekeverjük azt a tételt, hogy a mesterséges intelligencia hamarosan emberi képességeket szerezhet, azzal a tétellel, hogy a mesterséges intelligencia hamarosan az emberi természet részévé válhat - ez a hiba a tudományos fantasztikumban is megmaradt.<sup>18</sup>

---

Az erős mesterséges intelligenciát alátámasztó második állítás szerint a számítógépek hardveresen képesek lesznek utánozni az emberi gondolkodást. A számítógépek nyers teljesítménye a Moore-törvénynek megfelelően exponenciálisan növekszik. A Moore-törvény Gordon Moore-ról kapta a nevét,

- <sup>12</sup> *Lásd* HENRY BRIGHTON & HOWARD SELINA, INTRODUCING ARTIFICIAL INTELLIGENCE (422003).
- <sup>13</sup> *Lásd* Allen Newell & Herbert A. Simon, *Computer Science as Empirical Inquiry*: COMM19. OF THE ACM (113,1976118) ([link](#)).
- <sup>14</sup> John R. Searle, *Minds, Brains, and Programs*, BEHAV3. & BRAIN SCI. 417, 417-18 (1980).
- <sup>15</sup> *Lásd id.* 418
- <sup>16</sup> *Lásd id.*
- <sup>17</sup> *Lásd pl.* DANIEL C. DENNETT, CONSCIOUSNESS EXPLAINED (4391991).
- <sup>18</sup> *Lásd* a jegyzeteket 5257és- a kísérő szöveg.

az Intel egyik alapítója, az a megfigyelés, hogy a számítógépes chipekre illeszkedő tranzisztorok száma másfél-két évente megduplázódik.<sup>19</sup> Ez az előrejelzés, amely az elmúlt negyven évben megközelítőleg pontos volt, azt jelenti, hogy a digitális világ szinte minden aspektusa - a számítási teljesítménytől a számítógépes memóriáig - hasonlóan exponenciális ütemben növekszik.<sup>20</sup> A Moore-törvény tükrözi a számítógépeknek az emberiség alapvető motorjaként való gyors felemelkedését a huszadik század végén és a huszonegyedik század elején.<sup>21</sup>

Az exponenciális növekedés erejét nehéz túlbecsülni. Ahogy Robert Lucas mondta egyszer gazdasági vonatkozásban, ha egyszer elkezdünk az exponenciális növekedésről gondolkodni, "nehéz bármi másra gondolni".<sup>22</sup> Egy mobiltelefon számítási teljesítménye ma ezerszer nagyobb és egymilliószor olcsóbb, mint az MIT-n 1965-ben elhelyezett összes számítási teljesítmény.<sup>23</sup> Előre vetítve, harminc év múlva a számítógépek számítási teljesítménye valószínűleg egymilliószor erősebbnek bizonyul majd, mint a mai számítógépeké.<sup>24</sup>

Bizonyos, hogy néhány technológiai szakértő már régóta a Moore-törvény közelgő halálát jósolja, de ennek ellenére a Moore-törvény továbbra is virágzik. Az Intel, egy olyan vállalat, amelynek jelentős érdeke fűződik ahhoz, hogy pontosan megmondja a szoftvergyártóknak, mire számíthatnak,

---

<sup>19</sup> *Lásd* Moore törvénye: Made Real by Intel Innovation, <http://www.intel.com/technology/mooreslaw> (utolsó látogatás: 2010. március 29.) (a Moore-törvényt tárgyalja, amely szerint a szilíciumchipeken lévő tranzisztorok száma nagyjából 18-24 havonta megduplázódik, így a mikroprocesszorok sebessége rendszeresen növekszik) ([link](#)).

<sup>20</sup> *Lásd* Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, VA89. L. REV. n.1575,1620147 (2003).

<sup>21</sup> *Vö.* HENRY ADAMS, THE EDUCATION OF HENRY ADAMS 379-90 (1918) (a Szűzanyáról mint a középkor szimbólumáról és a gőzgépről mint a tizenkilencedik század szimbólumáról).

<sup>22</sup> Robert E. Lucas, Jr., *On the Mechanics of Economic Development*, J22. OF MONETARY ECON. 3, 5 (1988) ([link](#)).

<sup>23</sup> Ray Kurzweil, *Making the World a Billion Times Better*, WASHINGTONPOST.COM, 2008. április 13., <http://www.washingtonpost.com/wp-dyn/content/article/2008/04/11/AR2008041103326.html>

[\(link\)](#).

<sup>24</sup> *Lásd* HANS MORAVEC, ROBOT: 104-08 (1999).

előrejelzése szerint a Moore-törvény legalább 2029-ig fennmarad.<sup>25</sup> Ray Kurzweil technológiai teoretikus és feltaláló kimutatja, hogy a Moore-törvény valójában egy általánosabb exponenciális számítási növekedés része, amely már több mint száz éve erősödik.<sup>26</sup> Az integrált áramkörök felváltották a tranzisztorokat, amelyek korábban a vákuumcsöveket váltották fel, amelyek a maguk idejében az elektromechanikus számítási módszereket váltották fel. A számítási mechanizmusok mindezen változásai révén a számítás ereje exponenciális ütemben növekedett.<sup>27</sup> Ez a történelmi perspektíva azt sugallja, hogy a kutatás alatt álló új módszerek - a szén nanotechnológiától az optikai számításon át a kvantumszámításig - valószínűleg lehetővé teszik a számítási teljesítmény további exponenciális növekedését, még akkor is, amikor a szilíciumalapú számítás eléri fizikai korlátait.<sup>28</sup> Feltételezve, hogy a számítógépek számítási kapacitása továbbra is a Moore-törvénynek megfelelően növekszik, a számítógép hardverkapacitása valószínűleg eléri az emberi aggyal való egyenlőséget.

a és 20252030.<sup>29</sup>

Még ha ez az ütem nem is folytatódik, nehéznek tűnik, hogy

úgy vélik, hogy ezt a kapacitást nem fogják elérni az évszázad közepére.

A harmadik kérdés az, hogy a programozók képesek lesznek-e olyan szoftvereket biztosítani, amelyek a hardverben elért eredményeket a mesterséges intelligencia fejlődéséhez felhasználják. Kétségtelen, hogy ijesztő kihívások állnak előttünk, amikor olyan szoftvereket kell létrehoznunk, amelyek képesek megérteni az emberi gondolkodás által megragadott összetett valóságot. Egyesek szerint a számítási kapacitás korábbi növekedése ellenére a mesterséges intelligencia nagyrészt kudarcot vallott, és alig van mit felmutatnia.



---

<sup>25</sup> Jeremy Geelan, *Moore törvénye: "We See No End in Sight", says Intel's Pat Gelsinger*, SOA WORLD MAGAZIN (2008. május), elérhető a <http://java.sys-con.com/read/557154.htm> oldalon ([link](#)).

<sup>26</sup> RAY KURZWEIL, A SZINGULARITÁS KÖZEL VAN: AMIKOR AZ EMBER MEGHALADJA A BIOLÓGIÁT (672005).

<sup>27</sup> *Id.*

<sup>28</sup> A kvantumszámítás jó bevezetőjét lásd GEORGE JOHNSON, A SHORTCUT THROUGH TIME: AZ ÚT A KVANTUMSZÁMÍTÓGÉPHEZ (2003).

<sup>29</sup> *Lásd* KURZWEIL, *fenti* 26,125-27. o.

ötven év munkája.<sup>30</sup> Ez az értékelés túl szigorúnak tűnik. 1997 óta a számítógépek képesek legyőzni a világ legnagyobb sakkozóit.<sup>31</sup> A számítógépek által irányított autók képesek önállóan navigálni a városi forgalomban.<sup>32</sup> Ezeket a teljesítményeket aligha lehet az intelligens viselkedés erőteljes példáiként elvitatni. Persze igaz, hogy a sakk egy teljesen formális rendszer, és még az autóvezetésnek is korlátozott - bár kiszámíthatatlanabb - megoldandó problémakészlete van. De aligha meglepő, hogy a mesterséges intelligencia a formálisabb és kiszámíthatóbb környezetekben történő intelligencia létrehozásától az informálisabb és változékonyabb környezetek felé halad. Mindenesetre a szoftveres fejlődés a hardveres képességek növekedésével párhuzamosan folytatódik.<sup>33</sup>

Itt nem azt akarom bizonyítani, hogy az általános mesterséges intelligencia sikeresen lemásolja, majd felülmúlja az emberi intelligenciát, csupán azt akarom sugallni, hogy ez a lehetőség igencsak valószínű. Mindenesetre a hasznos mesterséges intelligencia terén nagyobb fejlődés várható. Még ha a mesterséges intelligencia nem is haladja meg az emberi intelligenciát, akkor is hasznos meglátásokat nyújthat a még megoldatlan problémákra. Az ilyen fejlődés nagyon hasznos lehet a társadalmi döntéshozatalban, mivel lehetővé teszi a számítógépek számára, hogy segítsék az embereket a társadalmi jelenségek magyarázatában és a társadalmi tendenciák pályájának és hatásainak előrejelzésében. 2020-ra a számítógépek várhatóan ellenőrizhető hipotéziseket állítanak fel; így nem kell majd

---

<sup>30</sup> Lásd BRIGHTON & SELINA, *Supra* note at 12, 23.

<sup>31</sup> Lásd IBM Research :DeepBlueOverview, <http://www.research.ibm.com/deepblue/watch/html/c.shtml> (Gary Kasparov sakkbajnok és a Deep Blue számítógép közötti mérkőzés felvétele) ([link](#)).

<sup>32</sup> Lásd: *CMU Robot Car First in DARPA Urban Challenge*, SPIE. ORG, Nov5,2007.,.

<http://spic.org/x17538.xml> (a DARPA kihívás városi környezetben zajló versenyének leírása és videója) ([link](#)).

<sup>33</sup> Lásd: *Data, Data Everywhere*, ECONOMIST.COM, 2010. február 25., [http://www.economist.com/specialreports/displaystory.cfm?story\\_id=15557443](http://www.economist.com/specialreports/displaystory.cfm?story_id=15557443) (utalva arra, hogy "a számítógépes alkalmazásokat vezérlő algoritmusok fejlődése évtizedek óta ugyanolyan fontos szerepet játszik, mint Moore törvénye") ([link](#)).



TÁRSADALOMTUDOMÁNY ALULRÓL FELFELÉ (1996). A társadalomtudósok már most is ilyen típusú adatváltozásokat végeznek a társadalmi hálózatok terén. *Lásd* Stephen P. Borgatti, *Identifying Sets of Key Players in a Social Network*, 12 COMPUTATIONAL & MATHEMATICAL ORG. THEORY 21, 33 (2006) (elismeri, hogy egy adott társadalmi hálózatban, ha a hálózatban lévő linkek 10%-a valójában nem létezik, az elemző véletlenszerűen 10%-kal variálhatja az adathalmazt, és újra alkalmazhatja a modellt a hálózat legfontosabb szereplőinek olyan halmazának azonosítására, amely "nem feltétlenül optimális a megfigyelt adathalmazra, de jó minőségű megoldást jelent a gráf egészének szomszédságára").

36 *Lásd* NASSIM NICHOLAS TALEB, *FOOLED BY RANDOMNESS: THE HIDDEN ROLE OF CHANCE IN LIFE* (A véletlen rejtett szerepe az életben).

*AND IN THE MARKETS*, at xliiii (az élet véletlenszerűségének alábecsülésére való hajlamot tárgyalva) (Random House. 2008) (2004).

a lehetséges megoldások értékelésével (beleértve az out-of-the-box ötleteket is), a mesterséges intelligencia segít a vészhelyzetekre való tervezésben.

## II. AZ AI FENYEGETÉSEI

Sok technológiaelméleti szakember nagyon optimista azzal kapcsolatban, hogy az erős mesterséges intelligencia képes lesz jelentősen javítani az emberi életet. A mesterséges intelligencia által működtetett robotok jelentősen megkönnyíthetik az életet, különösen a fogyatékkal élők és az idősek számára. A mesterséges intelligencia segíthet olyan felfedezésekben, amelyek meghosszabbítják az emberi élettartamot. Egyes technológiai teoretikusok, mint például Ray Kurzweil, szélsőségesen úgy vélik, hogy a mesterséges intelligencia akár ötven éven belül egyfajta technológiai utópiához vezethet, ahol az emberek nagyon magas jövedelemmel és nagyon hosszú, sőt talán korlátlan élettartammal rendelkeznek.<sup>37</sup>

Még azok is, akik attól tartanak, hogy az erős mesterséges intelligencia veszélyeztetheti az emberiség létét, félelmeiket arra alapozzák, hogy az ilyen mesterséges intelligencia lehetséges és viszonylag hamarosan bekövetkezik. Bill Joy, a Sun Microsystems korábbi vezető technológusa például nem vitatja Kurzweillel, hogy a példátlan technológiai felgyorsulás korába lépünk, amelyben a mesterséges intelligencia a jelenleginél sokkal erősebb lesz. De az ő kilátásai ezzel a fejlődéssel kapcsolatban mélységesen pesszimisták. Egy széles körben megvitatott cikkében "Miért nincs ránk szüksége a jövőnek" címmel, a következőket írja

riadót fúj, hogy az ember végső soron nem tudja irányítani ezeket a gépeket.<sup>38</sup>

A hatalom  
ereje

kritikája éppen abban rejlik, hogy elismeri a potenciális előnyök sokaságát.

---

<sup>37</sup> Lásd KURZWEIL, *fenti* 26,320-330122,. o.

<sup>38</sup> Lásd Bill Joy, *Why The Future Doesn't Need Us*, WIRED. COM, 2000. április, <http://www.wired.com/wired/archive/8.04/joy.html> ([link](#)). Aggodalmával nincs egyedül. Lásd KEVIN WARWICK, A GÉPEK MENETE: A MESTERSÉGES INTELLIGENCIA ÁTTÖRÉSE 280-303. (2004)("A gépek lesznek a domináns életformák a Földön.").

az erős mesterséges intelligenciától. Joy számára azonban, bármennyire is nagyok az AI előnyei, még nagyobb a kockázata annak, hogy elveszíti az ellenőrzést a létrehozott intelligencia felett. Úgy tűnik, hogy szerinte az ember a varázsló tanoncához hasonlít - túl gyenge és túl tudatlan ahhoz, hogy uralni tudja a mestergépeket. Joy álláspontja egy olyan sajátos félelem csúcspontját jelenti, amely a romantikus korszakra nyúlik vissza, és amelyet először Frankenstein szörnye képviselt, aki azt az elképzelést szimbolizálta, hogy "minden tudományos fejlődés valójában a pusztítás álcázott formája".<sup>39</sup>

A mesterséges intelligenciától való félelem a harctéren még közvetlenebb aggodalomra adhat okot a növekvő számítási teljesítmény. A nemzetek mindig is igyekeztek a technológiai innovációt felhasználni arra, hogy előnyökhöz jussanak a hadviselésben.<sup>40</sup> A számítástechnikai fejlődés napjainkban alapvető fontosságú a nemzetvédelem, vagy globálisabban fogalmazva a szuverén katonai verseny szempontjából. A Defense Advanced Research Projects Agency (Védelmi Korszerű Kutatási Projektek Ügynöksége) dollármilliárdokat költ egyre fejlettebb katonai mechanizmusok kifejlesztésére, amelyek egyre nagyobb számítási kapacitástól függenek.<sup>41</sup>

Nehéz túlbecsülni, hogy a robotika fejlődése, amelyet maga is a mesterséges intelligencia hajt, milyen mértékben alakítja át az Egyesült Államok hadseregét. Az afganisztáni és az iraki háború során egyre többféle pilóta nélküli légi járművet (UAV) használtak. Például tíz pilóta nélküli "Predator" 2001 volt használatban, és a

---

<sup>39</sup> Lásd RICHARD HOLMES, A CSODÁK KORSZAKÁJA: HOGY A ROMANTIKUS GENERÁCIÓ FELFEDEZTE A BOLYGÓSÁGOT.



A tudomány szépsége és rémsége n94. (Vintage Books 2010) (2008).

40 *Lásd* ROBERT FRIEDEL, A JAVÍTÁS KULTÚRA: TECHNOLÓGIA ÉS A NYUGATI

MILLENNIUM 85, 113, 131, 374 (2007) (a technikai innováció és a háború közötti kapcsolatáról).

<sup>41</sup> *Lásd* Noah Shachtman bejegyzését a Danger Room blogon, *DARPA Chief Speaks*, [http://blog.wired.com/defense/2007/02/tony\\_tether\\_has\\_1.html](http://blog.wired.com/defense/2007/02/tony_tether_has_1.html) (2007. február 20., 10:29) (a védelmi fejlesztésekről beszél, amelyek egy része a számítástechnikai innovációktól függ) ([link](#)).

2007 végén 180.<sup>42</sup> A jelentős számítási kapacitástól függő pilóta nélküli repülőgépek egyre fontosabb részét képezik a hadseregünknek, és 2020-ra a repülőgépek többségét alkothatják. <sup>43</sup> A robotok az ég alatt is fontos feladatokat látnak el, mint például az aknamentesítés.<sup>44</sup> Már fejlesztés alatt állnak olyan robotok, amelyek egyfajta különleges gyalogosként lézerrel bálnának a mesterlövészek megölésére összpontosítva.<sup>45</sup> Mások mentősökként fognak működni.<sup>46</sup> Nem túlzás azt jósolni, hogy húsz vagy huszonöt év múlva a háborút már túlnyomórészt robotok vívhatják. A mesterséges intelligencia által vezérelt csatatér másfajta félelmeket vet fel, mint amilyeneket a mesterséges intelligencia potenciális autonómiája kelt. Itt az az aggodalom, hogy az emberi rosszindulat miatt ezek az egyre jobb képességű gépek egyre nagyobb pusztítást és pusztítást fognak végezni.

### III. AZ AI LEMONDÁSÁNAK HIÁBAVALÓSÁGA ÉS A TILALOM TILALMA HARCTÉRI ROBOTOK

Joy a "lemondás" mellett érvel - vagyis az olyan technológiákról való lemondás mellett, amelyek erős mesterséges intelligenciához vezethetnek. Azok, akik aggódnak a mesterséges intelligencia technológiájának a harctéren való alkalmazása miatt, inkább a mesterséges intelligencia által működtetett fegyverekre összpontosítanak. De akár a lemondás, akár az új fegyverek korlátozásának célja, minden ilyen programot konkrét jogi tilalmakba kell átültetni. Ezek a tilalmak, legalábbis a jelenlegi technológia és a jelenlegi geopolitika mellett, minden bizonnyal

---

<sup>42</sup> Lásd P.W. SINGER, WIRED FOR WAR: (352009)21.

<sup>43</sup> Lásd Rowan Scarborough, *Unmanned Warfare*, WASH. TIMES, május 8., 2005, A1. o.

<sup>44</sup> Egy nagyszabású filmben népszerű beszámolót lásd: *The Hurt Locker* (First Light Production 2009).

<sup>45</sup> *Lásd* Singer, *Supra* note, at42, 111.

<sup>46</sup> *Id.* 112.

hatástalan. Így a nemzetek valószínűleg nem fognak egyoldalúan lemondani a számítási teljesítmény felgyorsítása mögött álló technológiáról vagy az e technológia további felgyorsítására irányuló kutatásról.

Ha az Egyesült Államok lemondana erről a technológiáról, az egész világ vesztes lenne. Az Egyesült Államok egyszerre virágzó kereskedelmi köztársaság, amely profitál a globális békéből és jólétből, és a világ hegemonja, amely képes biztosítani a globális béke és biztonság közjavát. Mivel a béke által biztosított jólétből más nemzeteknél nagyobb részt kap, mint más nemzetek, ösztönözve van a globális béke fenntartásának terheinek viselésére, amely nemcsak az Egyesült Államoknak, hanem a világ többi részének is előnyös.<sup>47</sup> A mesterséges intelligencia hatalmáról való lemondással az Egyesült Államok valójában nagyobb ösztönzést adna a szélhámos nemzeteknek annak kifejlesztésére.

Így az egyoldalú lemondás egyetlen reális alternatívája egy globális megállapodás lenne a mesterséges intelligencia által vezérelt fegyverek lemondásáról vagy szabályozásáról. Egy ilyen megállapodás azonban ugyanolyan leküzdhetetlen akadályokba ütközne, mint a nukleáris leszerelés. Amint azt az Iránnal és Észak-Koreával kapcsolatos közelmúltbeli események is mutatják,<sup>48</sup> nehéznek, ha nem lehetetlennek tűnik meggyőzni a gazember nemzeteket, hogy mondjanak le a nukleáris fegyverekről. Ezek a fegyverek nemcsak geopolitikai erőt és presztízst jelentenek az ilyen nemzetek számára, hanem az ilyen fegyverek előállításának és gyártásának bármilyen tilalmának ellenőrzése olyan feladat, amely meghaladja a nemzetközi intézmények képességeit.

A mesterséges intelligenciával kapcsolatos technológiák esetében az ellenőrzési problémák sokkal nagyobbak. Viszonylag kevés technológia épül

---

<sup>47</sup> *Lásd* John O. McGinnis & Ilya Somin, *Should International Law Be Part of Our Law?*, STAN59. L. REV. 1175, 1236-1238 (2007).

<sup>48</sup> *Lásd* Gustavo R. Zlauvinen, *Nuclear Non-Proliferation and Unique Issues of Compliance*, ILSA12 J. INT'L & COMP. L. (2006)593,,595

egy atombomba, de az erős mesterséges intelligencia eléréséhez sok út vezet, és még több olyan, amelyet valószínűleg még fel fognak fedezni. Ráadásul egy atombomba megépítéséhez jelentős infrastruktúra szükséges.<sup>49</sup> A mesterséges intelligencia kutatása egy garázsban is elvégezhető. Egy atombomba megépítéséhez nagyon jelentős erőforrásokra van szükség, amelyek meghaladják a legtöbb

a nemzetállamokon kívüli csoportok.<sup>50</sup>

A mesterséges intelligencia kutatását a következők végzik

a főiskoláknál nem gazdagabb intézmények, és talán még kevesebb jelentős forrásra lenne szükségük.

Joy elismeri ezeket a nehézségeket, de nem kínál plauzibilis megoldást. Sőt, a lemondás megvalósítására vonatkozó fő ötlete azt mutatja, hogy a célját lehetetlen elérni. Azt javasolja, hogy az informatikusok és mérnökök tegyenek egyfajta hippokratészi esküt, hogy nem vesznek részt olyan mesterséges intelligencia-kutatásban, amely potenciálisan a következő eredményekhez vezethet

A mesterséges intelligencia, amely képes kiszorítani az emberi fajt.<sup>51</sup>

De sok tudós valószínűleg elutasítaná

letenni az esküt, mert nem értenének egyet Joy előrejelzéseivel. Feltételezve, hogy néhányan letennék az esküt, sok kormány valószínűleg nem engedné meg tudósainak, hogy tiszteletben tartsák azt, mivel a számítástechnikai fejlesztések fontosak a nemzetvédelem szempontjából. Még ha a tudósok magukra is hagynák őket, valószínűleg figyelmen kívül hagynák az esküt, mivel a magánipar jelentős összegeket fizetne az ezen a területen elért eredményekért. Végül pedig a tudósok még akkor is nehezen tudnának eleget tenni egy ilyen irányelvnek, ha akarnának, mert nehéz lenne megjósolni, hogy milyen felfedezések fogják hosszú távon előrevinni a mesterséges intelligenciát.

Mindezen okok miatt egy globális lemondási szerződés, vagy akár egy, a mesterséges intelligenciával kapcsolatos fegyverfejlesztésre korlátozott szerződés ellenőrzése nem jöhet szóba. Valójában a relatív könnyű

<sup>49</sup> *Lásd* William J. Perry, *Proliferation on the Peninsula; Five North Korean Nuclear Crises*, ANNALS607 Am. Acad. Pol. & Soc. Sci. 78, 78-79 (2006) ([link](#)).

<sup>50</sup> *Lásd id.*

<sup>51</sup> *Lásd* Joy, *fenti* megjegyzés (link38).

a mesterséges intelligencia kutatásának végrehajtása azt sugallja, hogy - legalábbis a technológia jelenlegi szintjén - nehéz lenne egy nemzet számára a teljes mértékben a saját lakosai ellen irányuló mesterséges intelligencia kutatására vonatkozó ilyen tilalmat érvényesíteni. Még egy belföldi tilalom is ütközne az ilyen kutatások folytatásának jelentős ösztönzőivel, mivel az így létrejövő találmányok jövedelmező alkalmazásokat biztosíthatnak a területek szélesebb körében, mint a nukleáris fegyverek kutatása.

#### IV. FOGALMI HIBÁK AZ AI-TÓL VALÓ FÉLELMÉKBEN

Az erős mesterséges intelligenciától származó fenyegetések - mind az a félelem, hogy egzisztenciális fenyegetést jelent az emberiségre nézve, mind az a félelem, hogy a háborúkban nagyobb emberveszteséget fog okozni - eltúlzottak, mert fogalmi és empirikus zavarokon alapulnak.

##### A. *Az egzisztenciális fenyegetés*

Az egzisztenciális fenyegetés akkor oldható fel, ha jelentős lehetőség van arra, hogy

barátságos mesterséges intelligencia létrehozása.<sup>52</sup> A barátságos AI olyan mesterséges intelligencia, amely nem használja a

autonómia az emberiségre nézve veszélyessé válik. A barátságos mesterséges intelligencia melletti érvelés azzal kezdődik, hogy elvetjük azt az érvet, miszerint a fejlett mesterséges intelligencia szükségszerűen rendelkezik majd olyan akaraterővel, amely arra készítheti, hogy az emberiség helyébe lépjen. Az alapvető tévedés az



ilyen

---

<sup>52</sup> *Lásd: SINGULARITY INSTITUTE FOR ARTIFICIAL INTELLIGENCE, Supra note 4, 2. § (részletes programot kínál a barátságos mesterséges intelligenciához). Ebben az esszében elvonatkoztatok a barátságos mesterséges intelligencia megvalósításának részletes programjától. Ehelyett a barátságos mesterséges intelligenciát egy olyan célt kitűző, emberhez hasonló kognitív képességként definiálnám, amely nem dönt önállóan úgy, hogy kárt okoz az embereknek. Amint azt az alábbiakban tárgyaljuk, annak eldöntése, hogy milyen lépések növelik e cél megvalósításának valószínűségét, nehéz feladat lehet, és valószínűleg a legjobb, ha a Nemzeti Egészségügyi Intézetek által az orvostudomány területén jelenleg végzett támogatás-nyújtás keretében történik.*

A gondolkodásmódot az a tendencia jellemzi, hogy a mesterséges intelligenciát antropomorfizálják.<sup>53</sup> Az emberek, más állatokhoz hasonlóan, genetikailag arra vannak programozva, hogy saját (és rokonaik) jólétét sok esetben fontosabbnak tartásák, mint bármely más élőlény jólétét.<sup>54</sup> Ennek a motivációnak az oka azonban az evolúció történetében rejlik: azok az állatok, amelyek a saját jólétüket helyezték előtérbe, nagyobb valószínűséggel voltak sikeresek génjeik szétosztásában.

a következő generációk.<sup>55</sup> A mesterséges intelligencia nem lesz közvetlen terméke a biológiai evolúció, és nem is feltétlenül az ahhoz hasonló folyamatoké. Ezért hiba azt gondolni, hogy a mesterséges intelligencia szükségszerűen rendelkezik azokkal a túlságosan emberi tulajdonságokkal, amelyek a korlátok megkerülésére és a hatalom átvételére törekcszenek.

Ez nem jelenti azt, hogy nem lehet elképzelni erős mesterséges intelligenciát, amely képes rosszindulatúságra. A mesterséges intelligencia létrehozásának egyik módja lehet például az evolúciós folyamat bizonyos aspektusainak megismétlése, hogy a mesterséges intelligencia változatai más változatok legyőzésével fejlődjenek - ez egyfajta teremtési verseny. Azt gondolhatnánk, hogy egy ilyen folyamat nagyobb valószínűséggel szűlné egzisztenciális fenyegetéseket. Továbbá nem zárható ki, hogy a rosszindulat vagy legalábbis a hatalomra való törekvés a mesterséges intelligencia kutatásának egy bizonyos irányzatának emergens tulajdonsága lehet.

Sőt, még egy nem antropomorf emberi intelligencia is jelenthet fenyegetést az emberiségre, de valószínűleg kezelhető fenyegetést. A legnagyobb probléma

hogy az ilyen mesterséges intelligencia közömbös lehet az emberi jólét iránt.<sup>56</sup> Így a

például, ha csak nem másképp nem programozták, olyan módon oldhat meg problémákat, hogy

<sup>53</sup> *Lásd id.*

<sup>54</sup> *Lásd* ROBERT WRIGHT, A morális állat: evolúciós pszichológia és a mindennapi élet. 336-37 (1994). *Lásd még* John O. McGinnis, *Az emberi alkotmány és az alkotmányos jog: A Prolegomenon*, 78. CONTEMP. LEGAL ISSUES (211,1997213) (a gének megőrzésére és terjesztésére irányuló evolúciós törekvés leírása).

<sup>55</sup> WRIGHT, *fenti* megjegyzés 54,336-37. o.

<sup>56</sup> *Lásd: SINGULARITY INSTITUTE FOR ARTIFICIAL INTELLIGENCE (MŰVES INTELLIGENCIÉVEL FOGALMAZÓ INTÉZET)*, *fentebbi* lábjegyzet, 2. pont4. 3.2.3.

emberekkel szembeni károkozáshoz vezet. De a közömbösség, nem pedig a veleszületett rosszindulat sokkal könnyebben gyógyítható. A mesterséges intelligencia programozható úgy, hogy mérlegelje az emberi értékek a döntéshozatalban.<sup>57</sup> A kulcs az ilyen programozás biztosítása lesz.

A teljesen barátságos mesterséges intelligencia reális kilátása az első ok, amiért a kormánynak inkább támogatnia, mintsem szabályoznia kellene a mesterséges intelligenciát. Az ilyen támogatás erős ösztönzést nyújthat a mesterséges intelligencia fejlesztői számára, hogy komolyan vegyék ezt a kérdést. Ha az ilyen program révén kifejlesztett mesterséges intelligencia mechanizmusai megőrzik számítási előnyüket, akkor viszont hasznosak lehetnek abban, hogy megtalálják a másfajta mesterséges intelligenciákból eredő lehetséges veszélyek megelőzésének módjait. Az biztos, hogy ez a megközelítés nem biztos út a sikerhez, de sokkal gyümölcsözőbbnek és megvalósíthatóbbnak tűnik, mint a lemondás.

A barátságos mesterséges intelligencia támogatásának kérdése nagyon kényes kérdés. A kormánynak nincs elég tudása ahhoz, hogy egyértelmű követelményeket határozzon meg, amelyeknek egy barátságos mesterséges intelligencia projektnek meg kell felelnie. Nincs kellően világos definíciója annak sem, hogy milyen a barátságos mesterséges intelligencia végállapota. Ez a tudatlanság megakadályozhatja, hogy a barátságos mesterséges intelligencia eléréséhez díjat, vagy akár olyan köztes célt határozzanak meg, amely előrehaladást tesz lehetővé a barátságos mesterséges intelligencia elérése felé.

ez a végső cél.<sup>58</sup> A legjobb módja a támogatásnak az lehet, ha inkább kutatásként kezeljük.

projekt, mint például a Nemzeti Egészségügyi Intézetek által finanszírozott projektek.

A számítógépes és kognitív tudósokból álló szakértői bírálóbizottságok átnéznék a

projekteket, és kiválasztanák azokat, amelyek

<sup>57</sup> *Id.* § 2.

<sup>58</sup> *Lásd* Jonathan H. Adler, *Szembenézés a klímadíjjal: Rewarding Energy Innovation to Achieve Climate Stabilization*, 20-21 (Case Research Paper Series in Legal Studies, Working Paper No. 2010-15, 2010), *elérhető a* [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1576699](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1576699) oldalon (kritériumokat kínál arra vonatkozóan, hogy mikor működnek jobban a díjak, mint a támogatások) ([link](#)); William A. Masters & Benoit Delbecq, *Accelerating Innovation with Prize Rewards* 10 (International Food Policy Research Institute Discussion Paper No. 835, 2008), *elérhető a* <http://www.ifpri.org/sites/default/files/publications/ifpridp00835.pdf> *honlapon* (amely szerint a díjak hasznosak lehetnek a kívánt projektek ösztönzésére, ha a díjpályázat "megvalósítható, de nehéz célt [és] egyértelműen mérhető célt tűz ki, és a díjakat egy pártatlan hatóság kiszámítható módon osztja szét") ([link](#)).

célja egyrészt a mesterséges intelligencia fejlesztése, másrészt annak biztosítása, hogy az ilyen fejlesztéseket megfelelő biztosítékok kísérjék.<sup>59</sup> Eleinte egy ilyen programnak meglehetősen szerénynek és olcsónak kell lennie. Amint bebizonyosodik, hogy a program valóban előmozdítja a barátságos mesterséges intelligencia céljait, a programot ki lehetne terjeszteni.<sup>60</sup>

*B. A harctéri mesterséges intelligenciával kapcsolatos aggodalmak*

Nem mintha a mesterséges intelligencia hiányában a háborúk vagy a fegyverek megszűnének. A mesterséges intelligencia háborúra gyakorolt hatásairól úgy gondolkozhatunk, hogy végiggondoljuk, milyen következményekkel járna, ha a csatatéren az embereket technológiailag fejlett robotok helyettesítenék. Ez a helyettesítés legalább három szempontból valószínűleg előnyös lesz az emberek számára.

Először is, a robotok hatékonyabbá és kevésbé sebezhetővé teszik a hagyományos erőket bizonyos tömegpusztító fegyverekkel, például vegyi és biológiai fegyverekkel szemben. A világ kiegyensúlyozását, hogy az ilyen fegyverek - még ha csak kis mértékben is - kevésbé hatékonyak legyenek, előnynek kell tekinteni.

Másodszor, a hagyományos hadseregek többek között azért vetnek be halálos erőt, hogy megvédjék az emberi katonákat a haláltól vagy súlyos sérüléstől. Ha egy csatában csak robotok vannak veszélyben, egy nemzet nagyobb valószínűséggel alkalmaz nem halálos erőt, például kábító fegyvereket és a

---

<sup>59</sup> A szakértői értékelésnek a szabályozás helyett történő alkalmazása az olyan technikai kérdések kezelésére, amelyeket nem lehet bürokratikus megbízásokkal megoldani, olyasmiről, amit a szabályozási folyamat során általánosabban is meg kellene fontolni. *Lásd* Stephen P. Croley, *The Administrative Procedure Act and Regulatory Reform: A Reconciliation*, ADMIN10. L.J. (1996)35,47

<sup>60</sup> Ezért az itt megfogalmazott politikai ajánlások némileg eltérnek a Singularity Institute ajánlásaitól. Az ugyanis megvalósíthatatlannak tartja a lemondást. *Lásd* Creating Friendly AI, *supra* x. lábjegyzet, at Az intézet azonban 4.2.2. úgy tűnik, hogy a kormányzat részéről egy "hands off" megközelítést javasol. *Id.* at 4.2.1. Miközben egyetértek azzal, hogy a szabályozási tilalmak meggondolatlanok, a barátságos mesterséges intelligencia kormányzati támogatása hasznos, hogy lendületet adjon neki a potenciálisan kevésbé barátságos változatok feletti győzelemhez, amelyek lehetőségét szerintem nem lehet a priori kizárni. A mesterséges intelligencia potenciálisan barátságos, de nem feltétlenül az. Sőt, ahogyan azt az V. részben kifejttem, a mesterséges intelligencia pozitív externális hatásai a kollektív döntéshozatalban azt sugallják, hogy az ilyen támogatás indokolt.

mint. Az Egyesült Államok valójában azt fontolgatja, hogy egyes roboterejét nem halálos fegyverekkel szereli fel.

Harmadszor, a mesterséges intelligencia által vezérelt fegyverek előnyt jelentenek a fejlett világnak és különösen az Egyesült Államoknak, mivel a technológiai innováció terén fejlett képességekkel rendelkeznek. A robotizált fegyverek a legsikeresebbek közé tartoznak az Al-Kaida és más, az Egyesült Államok ellen aszimmetrikus hadviselést folytató csoportok elleni küzdelemben. A Predator, egy robotrepülőgép, sikeresen vette célba a terroristákat Afganisztán és Pakisztán egész területén, és a technológiailag fejlettebb változatokat gyorsan fejlesztik. Ráadásul mindezt célzottan teszi, anélkül, hogy nagyszabású háborúkat kellene indítani a területek megtartása érdekében - ez a folyamat majdnem

minden bizonnyal több járulékos kárt okozna.<sup>61</sup>

Ha valaki úgy véli, hogy az Egyesült Államok

összességében a békés és virágzó világot biztosító magatartási szabályok legjobb végrehajtója, ezt a fejlődést is előnyként kell értékelni.

Fontos, hogy a hadijogot hozzá lehet igazítani a robotok alkalmazásához. A polgári lakosság szándékos megkárosításának tilalmára vonatkozó alapvető követelmények nem lehetnek kevésbé érvényesek, ha a károkozók robotok. Hosszú távon a robotok, akár autonómok, akár nem, jobban képesek lehetnek a megkülönböztetésre, mint más típusú fegyverek, és így magasabb mércét alkalmazhatnak a polgári halálesetek elkerülésére.<sup>62</sup>

A háborúban az arányosság követelménye még inkább érvényesülhet, mivel az arányosság szükségessége

---

<sup>61</sup> Amint azt Kenneth Anderson professzor is kifejtette, lásd: *Why Targeted Killing? And Why Is Robotics so Crucial an Issue in Targeted Killing?*, <http://kennethandersonlawofwar.blogspot.com/2009/03/why-targeted-killing-and-why-is.html#links>



(március 1227,2009,;36) ([link](#)).

<sup>62</sup> *Vö. A drónok felemelkedése: Unmanned Systems and the Future of War: Hearing Before the Subcomm. on National Security and Foreign Affairs of the H. Comm. on Oversight and Government Reform 111th Cong. 11 (2010), elérhető a*

[http://oversight.house.gov/images/stories/subcommittees/NS\\_Subcommittee/3.23.10\\_Drones/Anderson](http://oversight.house.gov/images/stories/subcommittees/NS_Subcommittee/3.23.10_Drones/Anderson) oldalon.

.pdf (Kenneth Anderson írásbeli tanúvallomása) (a drónok és a tüzérségi sortűz által okozott járulékos károk összehasonlítása) ([link](#)).

a robotok minden sérüléstől való védelme kisebb lehet, mint az emberek védelmének szükségessége, így a nemzetközi jog által a csapatok védelmére ténylegesen engedélyezett erő arányosan kisebb lehet.

Így az Egyesült Államok lemondása súlyos hiba lenne, ha lényegében lehetséges lenne, hogy az emberi élet folyamatos virágzásával összhangban lévő mesterséges intelligenciát lehessen létrehozni.<sup>63</sup> Azt gondolhatnánk, hogy a mesterséges intelligencia bármilyen katonai célú kihasználása feszültségben áll a barátságos mesterséges intelligencia fejlesztésével. A számítógépek által irányított fegyverek által felszabadított pusztító erő azonban nem feltétlenül vonja maga után az erős mesterséges intelligencia létrehozását, amely az emberiség egészének kizorításához vezetne. Sőt, még ha a nemzetállamok közötti verseny végső soron nyomást gyakorolna is az erős mesterséges intelligencia katonai célú felhasználásának fejlesztésére, az Egyesült Államok számára erős ösztönzőket jelentene, hogy az ilyen fegyvereket működtető mesterséges intelligenciát visszatartsa attól, hogy olyan magatartást tanúsítson, amelytől Joy tart. Mindenesetre valószínű, hogy az Egyesült Államokban és más fejlett ipari nemzetekben más nemzeteknél jobban megbízhatunk abban, hogy figyelembe veszik ezeket a veszélyeket, különösen, ha folyamatban van egy barátságos mesterséges intelligencia programjuk. Így a baráti mesterséges intelligencia támogatásának és a számítógépes fegyverek kutatásának és az Egyesült Államok általi telepítésének kombinációja továbbra is jobb politika, mint a lemondás vagy a leszerelés alternatívája.

## V. AZ AI ELŐNYEI A FELGYORSULÓ TECHNOLÓGIA KORÁBAN

A felgyorsuló

technológiakorában élünk.

A Moore-törvény által szimbolizált és példázott exponenciális növekedésnek

köszönhető technológiai

innováció

---

<sup>63</sup> *Lásd: SINGULARITY INSTITUTE FOR ARTIFICIAL INTELLIGENCE, fentebbi lábjegyzet, §§4. 4.*

gyorsabban fejlődik, mint valaha. A technológia egyes teoretikusai úgy vélik, hogy az ilyen

a gyorsulás az emberi történelem során végigvonult.<sup>64</sup> Sokkal kevesebb időbe telt

hogy az ipari korszak felváltotta a földművelés korszakát, mint az, hogy ez a korszak felváltotta a vadászó-gyűjtögető vadászok hosszú évszázadát a távoli múltunkban. A posztindusztriális korszak még gyorsabban követi az iparosodást.

Még a mi életünkben is érezhető a változások felgyorsulása. Ha valaki ellátogat egy elektronikai boltba vagy akár egy élelmiszerboltba, két éven belül egy teljesen új termékcsaládot talál, míg valaki, aki 1910 és 1920 között - nem is beszélve 1810 és 1820 között - egy boltba látogatott, nem sok különbséget vett volna észre. Még a kulturális generációk is gyorsabban mozognak. A Facebook például néhány év alatt teljesen megváltoztatta az egyetemisták kapcsolatrendszerét,<sup>65</sup> míg 1940 a főiskolai élet hangulata felismerhető lett volna egy diák számára a 1970.<sup>66</sup>

A társadalom számára a legfontosabb, hogy a felgyorsuló technológia innovációk sokaságát hozza létre - olyan innovációs területeket is, amelyek három évtizeddel ezelőtt még ismeretlenek voltak. A biotechnológia például felveti az élet radikális meghosszabbításának lehetőségét, és ezzel együtt,

nagyon fontos változások a demográfiában. A nanotechnológia, amely a

az egy milliárdod méteresnél kisebb anyag irányítása és gyártása, szintén

---

<sup>64</sup> Lásd KURZWEIL, fenti megjegyzés, 26,17-19. pont.

<sup>65</sup> Maria Tess Shier, *The Way Technology Changes How We Do What We Do*, 2005 NEW DIRECTIONS FOR STUDENT SERVS. 77, 83 (2005) ([link](#)). Lásd még Matthew Robert Vanden Boogart, *Uncovering the Social Impacts of Facebook on a College Campus* (2006) (kiadatlan M.S. disszertáció, Kansas State University), elérhető a <http://krex.k-state.edu/dspace/bitstream/2097/181/1/MatthewVandenBoogart2006.pdf> ([link](#)).

<sup>66</sup> Ami a jogi folyóiratcikkek publikálását illeti, az elmúlt hat év során többet változott, mint az azt megelőző ötven évben. A Social Science Research Network (Társadalomtudományi Kutatási Hálózat) elterjedésével a tudományos cikkek vázlatai elérhetővé váltak - gyakran jóval a megjelenés előtt. A jogi szaklapok erre úgy reagáltak, hogy új módszereket találtak az értéknövelésre. A nagyobb jogi

folyóiratok például online komponensekkel egészítették ki a nyomtatott cikkeket, amelyek néha kommentálják azokat. *Lásd például* a Northwestern Law Review Colloquy-t, <http://northwestern-colloquy.typepad.com/> (utolsó látogatás 201012, áprilisában). ([link](#)).

<sup>67</sup> *Lásd* GLENN REYNOLDS, AN ARMY OF DAVIDS: HOW MARKETS AND TECHNOLOGY EMPOWER ORDINARY PEOPLE TO BEAT BIG MEDIA, BIG GOVERNMENT AND OTHER GOLIATHS 175-93 (2006).

széles körben folytatódik, és már beépült a termékekbe a naptejtol a növényvédő szerekig.<sup>68</sup> Bár a terület óriási ígéretekkel bír, potenciális veszélyeket is rejt magában. Ezek között vannak viszonylag hétköznapi veszélyek, mint például az apró részecskék által okozott testi károk, és sokkal vadabb forgatókönyvek, amelyekben a nanomasinák addig szaporodnak, amíg a Föld el nem pusztul.<sup>69</sup> Ezeken túl bizonyos területeken a felgyorsuló technológia valósága olyan teljesen új, bomlasztó innovációkat fog eredményezni, amelyeket ma még nem lehet előre jelezni. De ezek is gyorsan alakíthatják a társadalom szerkezetét jóra vagy rosszra.

Az ilyen technológiai változások néha új törvényeket, szabályozásokat és társadalmi normákat követelnek meg, hogy elkerüljük a veszélyeket és megvalósítsuk az ígéretek. Ezek a technológiák azonban önmagukban nem segítik elő közvetlenül a jobb információgyűjtést és -elemzést, amely e technológiai átalakulások felügyeletéhez szükséges. Más szavakkal, önmagukban nem biztosítják közvetlenül a gyorsuló technológiák szabályozását és társadalomba való integrálását elősegítő információkat. A mesterséges intelligencia ebben a tekintetben alapvetően más. Amennyiben a mesterséges intelligencia jótékony hatású marad, megkönnyíti az információk gyűjtését és elemzését, amelyek segítik a további fejlődés szabályozását nemcsak a saját területén, hanem a gyorsuló technológiák más területein is. További előnye, hogy a hangsúlyt az objektív információk feldolgozására helyezi. Bár a mesterséges intelligencia növekedése nem fogja megszüntetni az adatok értelmezésével kapcsolatos ideológiai csatározásokat, egyre több és objektívebben elemzett

---

<sup>68</sup> Lásd Diana M. Bowman & Graeme A. Hodge, *A Small Matter of Regulation: An International Review of Nanotechnology Regulation*, 8 COLUM. SCI. & TECH. L. REV. 1, 3 (2007),

<http://www.stlr.org/cite.cgi?volume=8&article=1> ([link](#)).

<sup>69</sup> *Lásd* Albert C. Lin, *Size Matters*: HARV31. ENVTL. L. REV. 349,355 (2007) ([link](#)).

a tények a mérlegelést segítik, megakadályozzák a szélsőséges és megalapozatlan állításokat, és a demokratikus konszenzushoz biztosítanak támpontot.

A mesterséges intelligencia fontosságának egyik megközelítése az, hogy a világban felgyorsuló információhalmazódás miatt egyre nagyobb szükség van az információk rendezésére és elemzésére szolgáló mechanizmusokra.<sup>70</sup> A hadseregnek már most is gondot okoz a drónoktól kapott összes információ elemzése, mivel nincs elegendő információ.

és elemzési kapacitás.<sup>71</sup> válogatási Ez a probléma a

a döntéshozatal egésze. Mivel a felgyorsuló technológia minden eddiginél gyorsabban teremt új komplexitást olyan területeken, mint a nanotechnológia, a biotechnológia és a robotika.

, a társadalmi döntéshozatalnak ugyanúgy meg kell küzdenie azzal, hogy lépést tartson az új jelenségek gazdagságának elemzésével, mint ahogy a hadseregnek is meg kell küzdenie a modern technológiától kapott egyre részletesebb információk feldolgozásával. A társadalmak akkor prosperálnak, ha a rendelkezésre álló összes információt fel tudják használni a lehető legjobb döntések meghozatalához. A probléma most az, hogy a feldolgozható információk meghaladhatják az emberi kapacitást, és a mesterséges intelligencia segítségével nélkül nem képesek megalapozott társadalmi döntéshozatalt elérni.

A nagyobb információfeldolgozási kapacitás egyik mellékhatása lehet, hogy a természeti katasztrófákat jobban meg tudjuk jósolni, és vagy megelőzzük őket, vagy megelőző intézkedéseket hozunk a legrosszabb következmények elkerülése érdekében. Minél kifinomultabbak a földrengések, az időjárás és az aszteroidák szimulációi és modellezései, és minél jobban össze lehet gyűjteni a



---

<sup>70</sup> *Lásd* ATUL GAWANDE, THE CHECKLIST MANIFESTO: HOW TO GET THINGS RIGHT (2010) (rajz figyelem a nagyobb információtenger által okozott problémákra). Gawande megoldása az orvostudomány területén az életmentés szempontjából legfontosabb protokollokra összpontosító ellenőrző listák készítése. *Id.* 137-38. o. Bár az ellenőrzőlistás megközelítés a jelenlegi technológia mellett valóban a legjobb megoldás lehet az információ túlterheltségre, az ellenőrzőlistával általában az a probléma, hogy egy olyan, mindenre egyformán alkalmazható megközelítést hoz létre, amely megreked az időben. Ezzel szemben a mesterséges intelligencia lehetővé tenné annak eldöntését, hogy mi a legfontosabb egy adott beteg számára, a naprakész információk figyelembevételével.

<sup>71</sup> *Lásd* Christopher Drew, *Military is Awash in Data from Drones*, NYTIMES.COM, Jan. <http://www.n10,ytimes.com/2010/01/11/business/11drone.html> (2010,[link](#)).

a jelenségekre vonatkozó hatalmas mennyiségű adat, annál jobbak lesznek az ilyen intézkedések.

legyen. A különböző katasztrófák kockázatainak becslésével a társadalom jobban képes en.<sup>72</sup> lesz arra, hogy

hogyan korlátozott erőforrásait a legsúlyosabbakra összpontosítva használja fel.

A technológia felgyorsulása páratlan előnyök kaszkádjait, valamint új katasztrófakockázatokat hozhat létre. Ez a felgyorsulás potenciálisan veszélyeztetheti az emberi faj jövőjét, de radikálisan meghosszabbíthatja az egyes emberek élettartamát is. Ha ez a felgyorsulás korunk alapvető jelensége, akkor a technológia következményeinek felmérése alapvető társadalmi feladat. Ebből következően a kormányzatnak különös érdeke fűződik annak a technológiának a felgyorsításához, amely elemezheti a technológiai gyorsulás többi részét - az AI-nak. Árnyalt és nehéz kérdés, hogy milyen mértékben és milyen formában indokolt ennek a technológiának a felgyorsulását fellendíteni, hogy segítse a többi felgyorsulással kapcsolatos döntéshozatalt. De ez a helyes kérdés, nem pedig az, hogy bonyolult szabályozásokkal hátráltassuk-e a fejlődését, vagy ami még rosszabb, mondjunk le róla.

---

<sup>72</sup> *Lásd* FLORIN DIACU, MEGADISZTEREK: A KÖVETKEZŐ KATASZTRÓFA ELŐREJELZÉSÉNEK TUDOMÁNYA (2010).