

Mesterséges intelligencia – a diskurzusok újratervésének kora

A közelmúltban az intelligens robotokban rejlő egyre növekvő veszélyek a mesterséges intelligenciáról (MI) szóló irodalom központi témájává váltak. Ez az általam „alarmistának” nevezett nézőpont logikus következménye (és erős szövetségese) az „erős MI” korosodó paradigmájának, illetve ezen paradigma legújabb változatainak. A veszélydiskurzus főbb argumentumainak és az ezekkel kapcsolatos problémák ismertetése után felvázolok egy új értelmezési keretet, melyben minden MI-rendszer elválaszthatatlan egységet, hibridet alkot humán komponenssel, a funkcióval és környezetével. Véleményem szerint az igazán fontos kutatási, tervezési és fejlesztési kérdések a humán összetevővel, valamint a humán és a mesterséges komponens interakciójával kapcsolatosak. Ez a megközelítés három további diskurzus számára is teret nyit: az automatizáció következő szintje és a foglalkoztatás jövője, az emberi összetevő kiterjesztése és az ehhez kapcsolódó jogi és etikai kérdések, melyek az MI és a robotika legújabb generációs fejlesztései vetnek fel.

Kulcsszavak: *alarmizmus, mesterséges intelligencia, hibrid rendszer, értelmezési keret*

Így hivatkozzon erre a cikkre:

Z. Karvalics László, „Mesterséges intelligencia – a diskurzusok újratervésének kora”.

Információs Társadalom XV, 4. szám (2015): 7–41.

<https://dx.doi.org/10.22503/inftars.XV.2015.4.1>

A folyóiratban közölt művek

a Creative Commons Nevezd meg! – Ne add el! – Így add tovább! 4.0

Nemzetközi Licenc feltételeinek megfelelően használhatók.

Z. Karvalics László

Mesterséges intelligencia – a diskurzusok újratervezésének kora

2015 szeptemberében számos népszerű magyar online portál adott hírt arról,¹ hogy a Queenslandi Műszaki Egyetem intelligens robotot készített az ökoszisztémában kulcsszerepet játszó korallzátonyok védelmében. Északkelet-Ausztrália partjainál ugyanis komoly részben a töviskoronás tengericsillag felelős a pusztításért, ezért a tudósok kamerával és GPS-szel víz alá merülő automata szerkezetet terveztek, olyan szoftverrel, amely fotók és videók ezrei alapján képes megbízhatóan azonosítani a kártevőt, a töviskoronás tengericsillagot, és sűrített levegőt befecskendezve végezni velük. Az elsőként üzembe helyezett masina 8 munkóra alatt mintegy 200 töviskoronás tengericsillagot öl meg, de a cél az, hogy a tapasztalatok alapján akár robotok százai is csatasorba álljanak, a minél nagyobb pusztítás érdekében.

Mivel az izgalmas fejlesztésről tudósító cikkek „robotgyilkosokkal”, „víz alatti mészárlásokkal” tálták a programot, rögtön megjelentek azok a kommentárok, amelyek a Terminátor-filmek hangulatát felidézve a „ma még tengericsillag, holnap ember?” kérdést tették fel, ügyet sem vetve arra, hogy a víz alatti robot gyakorlatilag a kapa funkcióját tölti be, amellyel egy ültetvényen gyomlálnak, csak nem emberi szem, izomerő és tüdő hozza mozgásba.

Ha egy szinttel magasabbra emelkedünk, akkor kiderül, hogy a beavatkozás szükségessége az ökológiai egyensúly felborulására vezethető vissza, de hogy pontosan mi okozta a felborulást, arról már megoszlanak a vélemények. Vannak, akik úgy vélik, hogy invazív fajról van szó, amelynek kiirtása kívánatos, mások szerint az ember felelős, mert megtrikította a tengericsillagok természetes ellensége, a gyönyörű házáért intenzíven halászott kagyló populációit.

Az elkeseredett viták két dolgot tettek egyértelművé: az okok és magyarázatok magas szintű világában a teljesebb kép kialakításához további tudások és ismeretek megszerzésére van szükség, miközben a megoldáskeresés azonnali cselekvést igényel. A probléma forrása azonban semmiképp nem a könnyen diabolizálható „gyilkos robot” környékén, hanem civilizációnknak a környezettel való viszonyát szabályozó meghatározottságok összetett világában keresendő.²

¹ A hír forrása: O’Callaghan (2015).

² Mindez még egyértelműbben kirajzolódik, mondjuk, a kínai nagyvárosok légszennyezésének csökkentését támogató mesterséges intelligencia-megoldások értékelésekor. Nem kérdéses, hogy a környezetszennyező, kontrollálatlan ipari tevékenység, a közlekedés, a fűtés és az időjárás változása okozzák a bajt, de az sem, hogy a megoldáskeresésben egyidejűleg fontos a kedvezőtlen hatások azonnali enyhítését lehetővé tevő beavatkozások és rövid távú döntések támogatása, valamint a problémák okait megszüntető vagy módosító, hosszabb távon érvényesülő, komplexebb szabályozási vagy kultúráváltási lépések megtétele, amelyeket szemléletváltásnak kell megelőznie. Az utóbbi igazi civilizációs kihívás, az előbbihez esélyt kínálhat egy olyan előrejelző rendszer, amely óriási adattömeg és számítási kapacitás birtokában napokkal a szennyezés koncentrációjának veszélyes megnövekedését megelőzően jelzi a közelgő gondot, és javaslatot is tesz annak orvoslására (üzemek időszakos bezárására vagy forgalomcsökkentésre).

http://www.technologyreview.com/news/540806/how-artificial-intelligence-can-fight-air-pollution-in-china/?utm_campaign=newsletters&utm_source=newsletter-weekly-computing&utm_medium=email&utm_content=20150903

Nagyjából ugyanekkor, 2015 júliusának első napjaiban, a Bécsben tartott információ-tudományi világkonferencia egyik (merthogy több volt) családias Global Brain szekciójának vitaindító előadásában Michael E. Arth, az UNICE (Universal Network of Intelligent Conscious Entities)³ alapítója a hallgatóságnak szegezte a kérdést: *emelve fel bátran a kezét az, aki szerint az erős mesterséges intelligencia nem jelent potenciális veszélyt az emberiségre nézve!* Arth csalódott arcát a mai napig látom magam előtt, merthogy az első sorokból volt módomban szemlélni az eseményeket. Nem azt várta, hogy tizenöt hallgatójából tizenhárom emeli majd fel a kezét. Egész előadását ugyanis arra kívánta felépíteni, hogyan készülünk már most fel a „gondolkodó gépek” által hordozott fenyegetésre.

S ez a hevenyészett szavazás jól jelzi ugyan, hogy a kutatók nagy része nem szükségszerűen kútmérgező publicisztikák badarságaitól befolyásoltatva alakítja ki saját, autonóm véleményét ebben a kérdésben – de ez mégis egy tudományos konferencia volt, amelyen Arth tudósként vett részt. Evvel meggyőzően illusztrálta is, hogy a pánikkeltéssel operáló bulvárnarratívák milyen erősen összecúsúznak a tudományosnak mondottakkal. S mivel ezt aggasztónak gondolom, vitaindítónak szánt írásom elején röviden ismertetem ennek a két oldalról támogatott veszélydiskurzusnak a fő argumentumait, szomorúan konstatálva, hogy *a friss könyvtermést szinte teljes egészében ide tartozó munkák* dominálják.⁴ Ezt követően bemutatok néhány, szívemnek kedves ellendiskurzust és szerzőt, majd utána javaslatot teszek egy alternatív megközelítésre. Mindvégig igyekszem az önállóként kutatható, érdeklődésre számot tartó témakörök és álláspontok nagy számát is érzékeltetni. Nem gondolom azonban, hogy a fentiekkel bármit sikerülne „megoldani”: leginkább példát, mintát szeretnék adni arra, miként lehetne kiszabadulni a jelenlegi tipikus gondolkodási keretrendszerekből, és másképp, máshol és más veszélyeket, illetve kritikus csomópontokat azonosítani, mint a pánik-irodalom. Roppant kíváncsi is vagyok, vajon ki, miért és mit oszt vagy utasít el ebből a megközelítésből, vagy milyen más gondolkodási útvonalat ajánl.

Musk, Hawking és más borzongatók

Elon Musk, a sikercég Tesla elnöke szerint *„minél fejlettebb lesz egy robot, annál kevésbé tiszteli majd a tervezőjét”*. Az önfejlesztő robotokkal felgyorsul a „szuperintelligens gép” kifejlesztéséhez vezető út, amelynek végén ez a masina *„spamnek nézheti, kiszűri és eltörli a Föld szférájáról az embereket”*, mint egy kéretlen üzenetet.⁵ *„Potenciálisan veszélyesebb, mint az atombomba”* (Pesthy, 2015). Nem kisebb tekintély véleménye mellett állt ki evvel, mint Stephen Hawking asztrofizikus, aki szintén osztja azt a nézetet, hogy a gondolkodó gépek a pusztá létünkre nézve is veszélyt jelentenek. *“Egy ilyen mesterséges intelligencia pillanatok alatt önállósítaná és folyamatosan, egyre gyorsuló tempóban újratervezné magát. Miközben mi em-*

³ <http://www.unice.info/unice/index.htm> (Letöltve: 2015. november 11.)

⁴ A kutatás első szakaszát a FuturICT.hu nevű, TÁMOP-4.2.2.C-11/1/KONV-2012-0013 azonosítószerű projekt támogatta az Európai Unió és az Európai Szociális Alap társfinanszírozása mellett. A befejezés *„Az Európai Unió és Magyarország támogatásával, a TÁMOP 4.2.1.D-15/1/KONV-2015-0006 azonosító számú - Ösztöndíj magyar és külföldi hallgatóknak és kutatóknak - A kőszegi innovációs kutatóbázis és tudásközpont fejlesztése a Pannon Egyetem oktatási és kutatási hálózatának keretében”* projekt részeként készült.

⁵ <http://www.origo.hu/techbazis/20141009-spamnek-nezik-es-torlik-az-emberiseget-a-robotok.html>
<http://www.vanityfair.com/online/daily/2014/10/elon-musk-artificial-intelligence-fear>

*berek, mivel fejlődésünknek határt szab a lassú biológiai evolúció, menthetetlenül lemaradunk, és végül kiszorulunk a versenytől. A gépagy végez velünk.*⁶

A Terminátor és a Mátrix fantáziavilágából a valóságba átemelt veszély-üzenethez⁷ képest óvatos figyelmeztetésnek tűnik az ezredforduló előtti időszak közhelyes fordulata, hogy a „gondolkodó gépek teljesítménye meghaladja az emberét”. Itt már evolúciós kontextusban bizonyul „fejlettebbnek” a mesterséges intelligencia a biológiaiánál.⁸ James Barrat könyvében a szuperintelligens gép az utolsó emberi innováció, amely egyúttal elhozza a „véget” (Barrat, 2014). Amint utoléri a mesterséges intelligencia az emberit,⁹ azonnal saját

⁶ <http://www.origo.hu/techbazis/20141203-hawking-szerint-a-gepagy-vegul-vegez-velunk.html>

⁷ „Az emberrel szemben cselekvő, öntudatra ébredt robot” képét a fikciós világban a 2001: Űrodüsszeia legendás HAL nevű gépe kezdte megalapozni, legfrissebb fejezetét pedig a 2014-es Ex Machina című film írta – a valóságban pedig az ipari robot-balesetek és a „gyilkos drónok” megjelenése erősítette fel. 2015 nyarán egy német autógyárban egy ipari robot okozott halálos sérülést, de nem „cselekvő intelligencia”, hanem „rosszkor és rossz helyen bekapcsolt gép” formájában. A krónika hiába jegyez fel számos ilyen eseményt (az elsőt 1979-ből, a Ford-gyárból), ezek valójában mind emberi hibára és nem a robot-mivoltra vezethetőek vissza. Ahogy a ma már thrillerekbe is beszivárgó kisebb, szűnyogszerű, méreggel gyilkoló, és nagyobb, golyózáporral támadó drónok sem maguktól ölnek, hanem az optikai modulon keresztül a potenciális áldozatot azonosító és döntést hozó földi irányítás révén. Az általunk teremtetett és életre keltett, de ellenünk forduló lényvel szembeni félelmek valójában kulturálisan és nem a technológia miatt kódoltak – véli a koblenzi egyetem két kutatója, Ulrike Barthelmess és Ulrich Furbach (<http://www.technologyreview.com/view/527336/do-we-need-asimovs-laws/>). S hiába az alapos és sok forrást használó újságírói munka, amellyel a cselekvővé váló robottal kapcsolatos sok kérdést megfelelő óvatossággal sikerül tárgyalni, ha a cikk címe ez: *Robot ölt embert, ez már az apokalipszis?* (Bolcsó, 2015). S hiába ’nem’ a válasz a kérdésre, már a leadben, tehát azonnal a kérdés után, a cím önálló életre kel: van, aki nem olvassa el az egész írást, csak a címet tudatosítja, a linkgyűjteményekbe is csak a cím kerül majd bele...

⁸ Nem is véletlen tán, hogy erős képvisellel jelent meg az iker-gondolat, hogy *a Földön kívüli élet keresése helyett is a Földön kívüli mesterséges szuperintelligencia keresése* felé kellene fordulni, hiszen a Kozmosz egészére lehet igaz a szintetikus élet(?)formák magasabbrendűsége. Az amerikai filozófus, Susan Schneider néhány NASA-közszereplő által is támogatott megközelítését részletesen bemutatja Stone (2014).

⁹ Néhányan a kritikus pontot a *'hazugságra képes robotban'* látják – sajnos makacsul azt a svájci kísérletet félremagyarázva, ahol az erőforrások megtalálását egymásnak jelző robotokat irányító „program-agyak” új generációi (hibridizált szoftverei) a források csökkenésekor „felülírták” a jelzés-parancsot, és nem értesítették a többi gépet, hanem maguknak tartották fenn az erőforrást. A kísérlet egy biológiai evolúciós helyzetet modellezett, mégis azóta a legtöbb hivatkozást a „nemcsak ölni, már hazudni is tudnak a robotok” szöveghelyzetben találjuk – úgy, hogy a „robot” nem „hazudott”, hanem a megadott kiinduló feltételek által szabályozott térben programot módosított, és ezt nem egy ember-robot, hanem előformált robot-robot játszmában tette (Fox, 2009). Hasonlóképpen takarhat számtalan különböző megközelítést a *'via-képes robot'* fejlesztésének programja (Hernandez, 2016). Ígéretes és fontos irány azt vizsgálni, hogy egy személyes ember-gép kommunikációs helyzetben miként „olajozhat meg” interakciókat, ha a gépi oldal az emberi humorra emlékeztető jegyeket is tud csatasorba állítani, ráadásul szituáció-érzékenyen és adekvát módon. Nem kevésbé érdekes kihívás szöveget generálni humoros karikatúrákhoz (<http://research.microsoft.com/en-us/um/people/horvitz/phumor.pdf>), amely egyúttal valódi ösvény a humor természetének a jobb megértéséhez is. De azt tűzni ki célul, hogy adott időn belül humorértő és humorgyártó mesterséges intelligenciánk legyen, illuzórikus, értelmetlen és erőforrás-pazarló. Az emberi intelligenciának egyrészt semmi szükségére nincs gépi humorpótlékra, jól elvan a maga kultúrába, közösségbe és cselekvésbe beágyazott vicc-ökoszisztémájával. Másrészt a humor (akárcsak a neologizmus vagy a metaforálás) magasrendű jelentésművelet (hiszen jellemzően több jelentés-sík keverésével és a jelentésvárakozások felülírásával fejeti ki hatását). Amikor még a sokkal alacsonyabb szintű jelentésműveletek világában is megoldatlan kihívások sorával kell szembenézni, mi indokolná egy erre ráépülő kutatási szint létjogosultságát?

túlélése lesz a számára fontos, és már csak azt a kérdést tehetjük fel, hogy megengedik-e majd nekünk, náluknál fejletlenebb lényeknek, hogy az árnyékukban létezessünk. Bostrom (2014) még csak a gorillákhoz hasonlítja az emberi fajt, de Brain (2015) nem viccel: friss könyvének címében a „második intelligens faj”, az érző és öntudatra ébredt mesterséges intelligencia számára az emberiség már olyan irreleváns, mint ma nekünk a csótány...¹⁰

Elég úgy feltenni a kérdést, hogy *a mesterséges intelligencia szolgálni vagy helyettesíteni fog-e minket* (ráadásul egy iskolásoknak szánt, több mint 200 oldalas népszerűsítő összefoglalóban, mint Del Monte, 2014), hogy világos legyen: az *alarmista* kiindulópont öngerjesztővé vált. ¹¹ Azt látjuk, hogy az „*emberit meghaladó mesterséges intelligencia*” létrejötté olyan *axiómává fejlődött, amelyre kétely nélkül épülnek elméletek és előrejelzések.*

A korábban erős mesterséges intelligenciának (Strong Artificial Intelligence, SAI), később mesterséges általános intelligenciának (Artificial General Intelligence, AGI) nevezett irányzat annak az elképzelt intelligens gépnek a megvalósítására szövetkezik, amelyik bármely intellektuális feladatot teljesít, amelyre az ember képes. Ehhez a célhoz ma¹² a *szingularitás-hipotézis* adja az üzemanycgot.

Legtöbbsen az irányzat egyes számú prófétájához, Ray Kurzweilhez és az ő emblematikus könyvéhez kötik a szingularitás, az emberit utoléró gépi intelligencia tézisének megszületését (Kurzweil, 2013/2005), pedig a diskurzus most kereken félszáz éves.¹³ Irving

¹⁰ Meg kell jegyeznünk, hogy a szuper-intelligencia okozta világvége-forgatókönyvek csak részei egy átfogóbb világvége-narratívának, amely az összeomláshoz vezető technológiai okok mellett a politikai és gazdasági szempontokra is súlyt helyez (Rees, 2004).

¹¹ A mikroelektromechanikai kutatások egyik „nagy öregje” az időutazással foglalkozó könyve után egy interjúban (Love, 2014) megad forgatókönyvet és határidőt is: 2045-re a legfejlettebb fajok (!) már nem emberiek lesznek. Egyébként az alarmizmus esetleges hasznáról és evvel összefüggésben az alarmistákkal való foglalkozás értelméről vagy értelmetlenségéről is folynak viták. Előkerül olykor a gazdaságtudományi analógia, ahol a stratégiai inflexiós ponthoz közeledő vállalatok életét meghatározó változásokat megérező és előre jelző kollégákra, az *úgynevezett kasszandrákra* nagyon is ajánlatos odafigyelni (Grove, 1998). Grove gondolatmenetével csak az a baj, hogy a „kasszandrák” családjának csak kis részére igaz, hogy ráéreznek a változásra, és megfelelően jelzik előre a jövőt. Az összes többi „kasszandrára” fordított idő és figyelem túlnyomó része nettó veszteség. Jól belátható mindez, ha arra gondolunk, vajon mennyit segít a világvége-jóslatok, a kozmikus katasztrófákat napra/óra pontosan előrejelző „prófécia” történeti és tipológiai elemzése a Földet fenyegető külső hatásokkal foglalkozó tudományos közösségnek, a dizasztrológusoknak. Az alarmista pozíció kialakításának és képviselésének súlyos tehertétele továbbá, hogy a jószándékú és meggyőződéses alarmisták mellett tömegesen foglalnak el pozíciót azok, akik szerint a figyelemgazdaságban sokkal eredményesebbek lehetnek alarmizmussal, mint mérsékelt, tárgyyszerű előrejelzésekkel. Ezért is kötöttünk ki az Allen Institute for Artificial Intelligence vezetője, Oren Etzioni által bevezetett és népszerűsített „Frankenstein-komplexum” helyett az alarmizmus kifejezés mellett: a művelődéstörténeti analógiánál fontosabbnak tűnik ugyanis, hogy az egymásra licitálás kultúrájára is utalhassunk. Ki mond riasztóbbat, rémisztőbbet a gépi intelligencia jövőjében rejlő veszélyekről? Mindennek azért van haszna is: a szélsőségek visszamenőleg teszik meztelenné a kevésbé radikális alarmizmus királyait is.

¹² Korábban elsősorban Hans Moravec hajlítgatta a vitateret, aki a mesterséges gépi faj megszületését 2030-2040 közé tette (Moravec, 1988), s tíz évre rá már a szuperintelligencia megszületésére épített (Moravec, 1998). A Moravecvel szembeni kritikák sokat finomítottak az erős mesterséges intelligencia híveinek álláspontján is.

¹³ Érdeklődőknek érdemes a különlegesen alapos, tanulmány-értékű Wikipedia-szócikkkel kezdeni a diskurzustörténeti ismerkedést: http://www.wikiwand.com/en/Technological_singularity.

John Good 1965-ben jelentette meg az „*első ultraintelligens gépről*” szóló tanulmányát, bevezetve az intelligencia-robbanás (*intelligence explosion*) tézisést is, amely az egyre okosabb gépek által gyártott még okosabb gépektől várja a fordulópontot (Good, 1965). Konceptiója egyetlen, korántsem lényegtelen apróságban különbözik a kortárs elméletektől. A Vernor Vinge és Ray Kurzweil által képviselt szingularitást az teszi elérhetővé, hogy a robbanás exponenciálisan gyorsul, és *elvezet* a szingularitáshoz. Good koncepciójában az ultraintelligens gép megteremtése *után* kezdődik a robbanás, amely az innovációt onnantól „átolja” a gépek térfelére. Az emberírta technikatörténet nála itt véget is ér.

Belátható, hogy az alarmista nézőpont csakis azért született meg, hogy lelkiismereti ellensúlyt képezzen, és alternatív forgatókönyvet adjon a magasra emelkedett elvárás-horizontnak avval kapcsolatban, hogy mi várható néhány évtized múlva, az évszázad közepére¹⁴ a szingularitástól az emberiségre nézve. Kurzweilék, akiket emiatt utópistáknak tartanak, ugyanis nem aprózzák el az ígéretek: az üzlet világában kreatív rombolás megy végbe, nő a gazdagság, csökken az egyenlőtlenség, a világproblémák megoldhatóvá válnak, a biológiai tökéletesítéstől fantasztikus egészség és halhatatlanság-közelség várható (Miller, 2012). A szuperintelligenciához hasonlóan ugyanis axiómává vált az emberi agynak a minőségi agy-számítógép kapcsolathoz szükséges feltérképezése, megértése (Kurzweil, 2013), ami a szingularitással elért végállapot idején az agy „feltöltését”, áttöltését teszi majd lehetővé a biológiai hordozóról a gépre. Blackford és Broderick (2014) már ennek a jövőjéről, Rothblatt (2014) pedig az ekképpen elért „virtuális emberi” és a „digitális halhatatlanság” ígérteréről és veszélyeiről írt könyvet.

Szívesen nevezem végül *navigacionistáknak* azokat a szellemi műhelyeket és szerzőket, akiknek se a gyilkos robot rémképére, se a digitális halhatatlanságra nincs szükségük ahhoz, hogy az „intelligenciarobbanás” (*intelligence explosion*) következetesen a 21. század közepére várt pillanatához igazítsák óráikat. Ahogy Muehlhauser (2013) fogalmaz: történelmünk legfontosabb pillanata lesz az, amikor a gépi intelligencia és az avval összekapcsolt képességek meghaladják az emberit – s ennek a folyamatnak a bölcs navigálása a jövő leglényegesebb kihívása. Ha jól nyúlunk bele az eseményekbe, a szuperintelligencia a legjobb dolgok egyike lehet, ami az emberiséggel valaha történt, mert amihez nem vagyunk elég okosak, „ő”, a barátságos mesterséges intelligencia (*Friendly AI*)¹⁵ majd megoldja (Bostrom, 2014). Ellenben minden félre is csúszhat, ha nem vagyunk elég felkészültek. Ez a felemelt mutatóujj, ami a szélsőségekbe forduló alarmistákhoz képest tárgyilagosságot és „középutasságot” ígér, dollármilliókat ér a navigacionistáknak. Egyre-másra születnek azok a kutatóintézetek, amelyek a szingularitással foglalkozó kutatók legújabb generációját vonzzák be „a jövő megalkotásába”, miközben komoly forrástömeget tudnak mozgósítani.¹⁶

¹⁴ Az időpontról nagy (és jórészt értelmetlen) vita folyik, számtalan variáció kering, volt már, aki „átlagot” is számolt (2040-et), de a mindenki közül a leginkább radikális Goertzel (2014) szerint a szingularitáshoz akár jóval korábban is eljuthatunk, ha projektként megfelelő figyelem övezi és kellő mennyiségű forráshoz jut. Nem véletlenül nevezi Daniel C. Dennett városi legendának a szingularitás-hipotézist (Brockman, 2015).

¹⁵ A barátságos AI fogalmát (talán Casper, a barátságos szellem mintájára) Eliezer Yudkowsky alkotta meg, az alarmistákat ellensúlyozandó. Levy (2008) azt mutatta meg, *miért épülhet ki szoros érzelmi kapcsolat, valódi kötődésen alapuló szűzetség* (real companionship) ember és állat, *ember és robot között*, s hogy ez miért nem jelenti a „humán” dimenzió elvesztését.

¹⁶ A 2000-ban alapított Szingularitás Intézet (*Singularity Institute for Artificial Intelligence*) nőtt át a gépi intelligencia-kutató intézetté (*Machine Intelligence Research Institute*, MIRI). A 2005-ben az Oxford Egyetemen alapított *Future of Humanity Institute* (FHI) a filozófiai oldal művelésére jött létre, a rivális Cambridge 2012-ben hozta létre a maga központját (*Centre for the Study of Existential Risk*). A 2014 óta működő *Allen Institute for Artificial Intelligence* (<http://allenai.org/>) a Microsoft-alapító Paul Allen nevét viseli, s egyetemekkel közösen valósít meg (ebben a pillanatban) négy kutatási programot.

A magam részéről zsákcúcnak és már a fogalmi alapvetés oldaláról is elhibázottnak tartom az erős mesterséges intelligencia programjából kinövő szuperintelligencia (*Artificial Superintelligence, ASI*) fogalmát,¹⁷ és emiatt szükségszerűen értelmetlennek minden ráépülő narratívát.¹⁸ A szuperintelligencia tétele és az alarmizmus találkozása ugyanis a diskurzusok olyan terét teremti meg, amely feltételezett fejlődési folyamatokból adott eséllyel következő, lehetséges jövőállapotok tipikusan negatív kimenetei felé tereli a párbeszédet, miközben a mesterséges intelligenciának számos működő és éppen bevezetés előtt álló alkalmazása és megvalósulása van, amelyekkel kapcsolatban egészen más típusú kérdések vetődnek fel. Természetesen ott, ahol a jelfeldolgozás összekapcsolódik műveletvégzéssel, ahol az automatizáció átrendezi a folyamatok feletti kontroll-láncokat, és újra és újra más szerepet ad a gépnek és az embernek, kérdések garmadája merül fel, amelynek esetében indokolt a lehetséges veszélyek számbavétele is. A veszélydiskurzusok elsődleges szintjét azonban a létező megoldásokban és alkalmazásokban fellelhető vagy elképzelhető hibaforrások azonosítása és tárgyalása jelenti, erős kitekintéssel az érlelődő, készülődő, bevezetés előtt álló rendszerek működése során felmerülni vélt problémátípusokkal való mérnöki szembesüléssel. S miközben a szellemi közélet tipikusan az utópisztikus megközelítéseket utasítja el, sokkal több nehézséget okoz a gondolatrest disztópia, a veszély kiterjesztett, absztrakt, ismeretelméleti-filozófiai dimenzióinak összekeverése vagy összezsúvátása a valóságos kihívásokkal.

De magára a pánikkeltésre is egyáltalán azért kerülhet sor, mert a mesterséges intelligenciában rejlő fejlődési és veszélypotenciál minden csatornán mértéktelenül eltúlzottan jelenik meg, az akadémiai publikációktól az ipari küldetésnyilatkozatokon át a fikciós irodalomig. Ugyan milyen világvége-veszélyt idézne fel egy minden eddiginél sikeresebb nyelv- vagy arcfelismerő program? A mesterséges intelligencia-alkalmazások világát össze kell boronálni a robotokéval, amelyek ebben a pillanatban banális (és kifejezetten „buta”) mechanikai művelet-végző szerkezetek. A jelen valóságos robotjai a legkevésbé sem intelligensek, a legintelligensebb alkalmazások viszont nem lépnek át a szimbólumok teréből a valóságos fizikai terekbe, nincsenek „végrehajtó” kimeneteik, eredményeik emberi tevékenységek bemeneteit támogatják.

¹⁷ Természetesen sokkalkal együtt vélem így. Legutóbb Robert Trappl, az Osztrák Mesterséges Intelligencia Kutatóintézet (OFAI) igazgatója szögezte le: „A mesterséges intelligenciának bizonyos dolgokat jobban kell csinálnia az embereknél, egy szuperintelligencia kialakítását viszont nagyon valószínűtlennek tartom” <https://sg.hu/cikkek/114196/a-programok-tobbet-tudhatnak-majd-az-embereknel> (Letöltve: 2015. december 3.). S ahogy az Isten-érvek és Isten-ellenérvek története is megírható és feltárható, ugyanígy érdemes lehet egyszer végigszemlélni a gondolkodó gép mellett és ellen szóló érveket. Érdeklődőknek kiindulásképp javaslom Rovenszkij, Ujemov és Ujemova több mint félszáz éves, magyarul is megjelent könyvének utolsó fejezetét – *Miért nem lehetséges az elektronikus agy?* (Rovenszkij és társai, 1964, 206-216. o.).

¹⁸ Elsősorban azt a komikus igyekezetet, ahogyan a szuperintelligencia létrejöttének időpontját próbálják valamilyen nevezetes évhez kötni. Annak, hogy 2020 és 2060 között mikor valósul meg az „átfordulás”, ma már önálló könyvszerte van. De ugyanígy értelmetlen halmaznak tűnnek azok a „receptkönyvek” is, amelyek listába gyűjtik, „mindössze” mire van szükség a szuperintelligenciához: nagyobb számításteljesítményre (!), az agy működésének maradéktalan megértésére és szimulálhatóságára, az evolúciós szcéna mesterséges előidézésére és a gépi én-tudat megteremtésére (Urban, 2015). Mindez elvileg és még inkább a „mesterséges élet” és a mesterséges „komplex élő rendszerek” programjától lehetne remélni, de ebbe az irányba sokkal kevesebben tapogatóznak.

De mi a baj az erős mesterséges intelligenciával?¹⁹

Az erős mesterséges intelligencia ellenfelei hagyományosan azzal érvelnek, hogy még a „józan ész” egyszerű kihívásait²⁰ megérteni és kezelni tudó masinák létrehozatala is teljesíthetetlen kihívás: a könnyen algoritmizálhatónak tűnő, valójában szinte végtelen számú függő háttérváltozó ismeretét és használatát igénylő banális kommunikációs helyzetek és élethelyzetek modellezésébe is rendre beletörik a géppel szimulált elmék bicskájá. A „józan ész” alapján okoskodásra, párbeszédre és tanulásra képes gépek fejlesztésével kapcsolatban két irány bontakozik ki (Havasi, 2014). Az egyik még inkább logikai útvonalakra emlékeztető, szabály-alapú reprezentációkra tenné képesebbé az okos masinákat, a másik inkább még asszociatívabb, még inkább analógia-alapú, de a természetes nyelvre, gondolkodásra és viselkedésre jobban emlékeztető képességekkel ruházná fel őket.²¹ Ezek az irányok is arra épülnek azonban, hogy a mesterséges intelligencia-rendszereken belül információs folyamatok zajlanak.

Valójában azonban egyetlen *jel-feldolgozásban verhetetlen gép sem intelligens*. Ahogyan azt John R. Searle a nyolcvanas években már meggyőzően kifejtette, a kimenetnek mi adunk és tulajdonítunk értelmet, a feltételezett jelentés komplexitása vagy autenticitása miatt látunk bele intelligenciát. A gépben nem zajlik információfeldolgozás, csak kódmanipuláció. A gép jel-műveleteket hajt végre, programjának megfelelően: a komputáció lé-

¹⁹ Az (erős) mesterséges intelligencia dekonstruálása egyidős annak programjával. Alapvetően két ellenérv-típus jelenik meg ennek irodalmában: a *külső* (ontológiai-ismeretelméleti) és a *belső* (a technológiai korlátokra, költségekre és limitációkra figyelmeztető). Mivel ezek kézikönyv- és bevezető kurzus-szerűen jól ismert diskurzusok, csak olyan mértékben utalunk rájuk, amelyre szükség van a „szuperintelligencia”-narratíva megkérdőjelezéséhez. Az úgynevezett „erős mesterséges intelligencia” programjával szembeni érveket legújabbán összefoglalja Gleiser (2014). Igyekszünk azokat a szempontokat összegyűjteni, amelyek a viták legújabb hullámában merültek fel, s amelyek saját álláspontunk kifejtését segítik.

²⁰ Feladni egy menü-rendelést a vendéglőben? Magyarázkodni az ellenőrnek a jegy-lyukasztás késedelme miatt? A megfelelő zoknit kiválasztani az alkalomhoz, a cipőhöz és a ruhához? Más ruhában, más hajjal, öregebben, de ugyanaz az ember van-e a fényképen? Hogyan működhet a gépi tanulás ilyen helyzetekben? Megtanítható-e a gépek következő generációja a természetes nyelv feldolgozásának (natural language processing, NLP) magasabb szintjére, hogy kiállja a „józan ész” próbáit? A palacsintakészítő robot (és mögötte a brémai RoboHow projekt) azzal kísérletezik, hogy az emberi instrukció és az emberi viselkedés megfigyelése és adaptálása pótolja a „gépi” józan ész. Evvel viszont valójában nyelvfelismerési irányba viszik el a kutatást, miközben elvileg arra kellene képessé tenni a kísérlet ember-robot csapatát, hogy egy újfajta tejesdoboz kinyitására és megfelelő dőlésszögű kiöntésének módjára személyes instrukciókkal vezesse rá az ember a gépet. (Ehhez képest a palacsintareceptek letöltése a WikiHow oldalról banális részfeladat. <http://444.hu/2015/08/25/internetrol-tanul-palacsintat-kesziteni-egy-nemet-robot/>)

²¹ A logikai irány legismertebb sikertörténete Watsoné, az IBM műveltségi vetélkedőt nyerő robotjác (kevesebben tudják, hogy a Cycorp mai napig működő, Doug Lenat nevéhez kapcsolódó Cyc projektje 1984 (!) óta fejleszti azt a rendszert, amely a józan ész világának tényeit (újabb kezdeményezésekkel párhuzamosan) egyre nagyobb repozitóriumokba rendezi. A nyelvi alapú rendszerek legismertebbike az 1999-es Open Mind Common Sense kezdeményezés. Az egyik legkorábbi online önkéntességre (crowdsourcing) alapuló projekt eredménye a ConceptNet, amelynek szöveges információbázisa és hatalmas tudásgráfja 17 millió tényt tartalmaz különböző nyelveken.

nyege a számolásteljesítmény. „Tudása” a műveletkéesség, de nincs a műveletre magára vonatkozó „metaszintje” – mint amikor valaki elsajátítja, hogyan kell összeadni, kivonni, szorozni, osztani, de nem tudja, miért, mikor, minek az érdekében van szükség minderre. Ha mondják, és megkapja a számokat meg a műveleti utasítást, elvégzi, de maga soha nem kezdeményez, hiszen azt sem tudja, hogy alkalomadtán az egyes számok milyen mennyiségeket reprezentálnak, és mi felel meg nekik a valóságban.

Ezért is hívja a rendszerszemléletű információtudomány karizmatikus alakja, Alva Noë vagy az adattudós-jövőkutató Ken Bodnar a mai mesterséges intelligenciát *pszeudo-intelligenciának*. A legfejlettebbnek vélt programok, amely bizonyos teljesítményt (például emberi érzelmek rendkívül bonyolultnak tartott felismerését és azonosítását) lehetővé tehetik, egydimenziósak: nincsenek alárendelt princípiumaik, amelyek értelmet adnak a műveletnek, vagy a „miért” kérdésre válaszolni tudnának. Az imponálóan tűnő kimeneti oldal ellenére ezért ez mégiscsak a legostobább intelligencia (dumbest intelligence), hogy ismét csak Bodnar (2015a) idézzük.

Ken Goldberg a budapesti Brain Bar rendezvényén legutóbb az egyéves gyerek észlelési és manőverezési szintjére való felfejlesztést is rendkívül nehéz, óriási számítási erőforrást igénylő feladatnak nevezte. Yann LeCun az egér szintjét tartja nem meghaladhatónak, de Noë szerint még egy amőba is többet tud, mint a legintelligensebbnek mondott masinák. „Az egyetlen sejtnek élettörténete van; környezetét alakítja át azt a médiumot, amelyben találja magát, és ezt a környezetet értékes helyé szervezi. Tápanyagot keres. Megcsinálja magát – és azzal, hogy megcsinálja magát, értelmet visz az univerzumba.” A géppel ellentétben „az amőbának ... van információja – begyűjti és feldolgozza azt”.²² Akkor kezdünk majd aggódni a szingularitás miatt, amikor az IBM olyan gépeket gyárt, amely egy amőba működését és tudatosságát produkálja.²³

²² Noë gondolatait remek összefoglalásban idézi Pesthy (2015). Érdekes módon az amúgy néha bizarr teológiai kiindulópontú információs elméletek is hasonló következtetésekre jutnak „... a mesterséges intelligencia programok – bármily komplexnek és „intelligensnek” tűnnek is számunkra – csupán reprodukált, tehát semmiképpen sem kreatív információt jelentenek. A reprodukált információ létrehozásához nincs szükség saját szellemi tevékenységre, ezért ez a munka rábízható a számítógépekre” (Gitt, 2004, 164. o.).

²³ Mindez annak ismeretében erős kontraszt, hogy amióta a japán To-Robo szoftver magasabb pontszámot kapott egy főiskolai felvételi angol nyelvi részén, mint a japán diákok átlaga, azóta a világsajtót elárasztotta a szenzációs hír, hogy a mesterséges intelligencia teljesítménye már felülmúlja a főiskolásokét (<http://blogs.wsj.com/japanrealttime/2014/11/04/artificial-intelligence-outperforms-average-japanese-high-school-senior/>). Az amőba egyébként nem a valóságtól elrugaszkodott választás: miután egy alig félezer génnel rendelkező baktériumot már sikerült számítógéppel modellezni, az Open Worm projektnek (<http://www.openworm.org/>) egy sokkal bonyolultabb lény, a húszszor annyi génnel rendelkező fonálféreg (*Caenorhabditis elegans*), az első digitális élőlény megalkotása a célja, amelynek mozgása és viselkedése megfeleltethető a valóságos lénynek. Hasonló vágy fűti az első robot-orangután megalkotóját, Steve Grandet, aki virtuális környezetben kísérletezik valódi életforma létrehozásával. Terepményeit, a Grandroidokat ne úgy képzeljük el, mint egy játék figuráit: virtuális neuronokkal, receptorokkal, enzimekkel és génekkel ellátott ágensek ők. Polipszájjal rendelkező, borotvált kutyára emlékeztető lények, akik *viselkedni* tanulnak, és környezetükkel úgy tartanak kapcsolatot, hogy az interakciókról maguk „döntenek” (Parkin, 2015). De ha egy Grandroid felnő, vagy ha a mesterséges intelligencia el is jut a fonálféreg szintjéig, hol volna a hároméves gyermektől? A kérdést Alison Gopnik, a gyermeki intelligencia kutatója teszi fel, annak a kétségének adva hangot, hogy a gép intelligencia valaha is eljuthat-e egy kisgyermek szintjére. George Dyson ehhez azt az érvet illeszti, hogy a valódi kreatív gondolkodás mindig analóg marad, soha nem válhat digitálissá (Brockman, 2015).

S ezért jellemzi egyúttal az *autonómia hiánya is* a mégoly erősnek és okosnak mondott masinákat is. Nincsenek céljai, nincs akarata, nincsenek referenciapontjai, amelyhez viszonyítva a magára, a környezetre és a már korábban kialakult jelentésekre való tekintettel kellene új jelentéseket létrehoznia, és annak alapján döntést hoznia.²⁴ Programot futtat le, de az új program teremtése is program, és mindig hiányzik mögüle a jelentés. Minden, szemantikusnak mondott rendszer a jelentéssel mint speciális, gépi nyelven leírható, kódolhatóvá tett objektummal foglalkozik: érzékeny rá, de nem „érti”, nem teremti, nem használja, hanem a számára értelmezhető jellé lefordított formájában komputálja. Az intelligens rendszer nem egyszerűen létrehozza, módosítja, folyamatosan átalakuló alakzatokba szervezi a jelentéseket, hanem minden egyes észlelési ciklus kezdetén újratemti, újraértelmezi azokat, és azonnal összekapcsolja a jövő valamilyen előre jelzett képéhez igazodó cselekvés tervezésével. Információja nem létezik önmagában, hanem egy időtengely mentén felépülő, csakis történetiségében értelmezhető, folyamatosan változó, komplex kognitív univerzum részeként, mozgásában. A gépnyelvre lefordított információ statikus, és a kognitív univerzum valamely szakaszának vagy elemének képesség- vagy kapacitás-korlátját segít lebontani.

A biológiaiilag adott információkezelő képesség meghaladása már régóta zajlik extraszomatikus információtechnológiai megoldások fejlesztésével. Amikor az optikai eszközök segítségével az észlelés tartományait a makro- és mikrotartományokban kiterjesztjük, senki nem nevezné „intelligensebbnek” a távcsövet vagy a mikroszkópot az emberi szemnél. Csak nagyobb felbontásúnak. Amikor memóriánk korlátait listákba és katalógusokba tárgyiasított módon küzdjük le, ezeket a dokumentumtípusokat sem mondjuk „intelligensebbnek”, mint az embert, hanem egyenesen az emberi intelligencia meghosszabbításaként tekintünk rájuk. Amíg a számológép nem automatikusan, hanem emberi beavatkozást igényelve növeli meg a számolásteljesítményt, addig fel sem vetődik, hogy intelligens volna a masina, amely segít a komputációban. S amíg az immár automata gép csak számol és adatot dolgoz fel, addig sem tűnik intelligensnek – csak akkortól kezdve, amikor eltűnik a szemünk elől a műveletvégzés folyamata és logikája, és csak annak, számunkra jelentést hordozó kimenetét látjuk.²⁵

Általánosítva: az már régóta nem kérdés, hogy a számítógép teljesítménye a számolásműveletekben és a számolásművelet-alapú másodlagos műveletekben (mint például az alakfelismerés vagy a jel-visszakeresés) megfelelően formalizálható körülmények között felülmúlja az emberi teljesítményt, és a technológiai fejlődéssel párhuzamosan a teljesítménykülönbség mértéke egyre nő. (Mint ahogy az is teljesen elfogadott kiindulópont, hogy az algoritmizálható agymunka gépesíthetősége minden esetben a magasabb értékhozzáadású szellemi műveletekhez szabadtítja fel az emberi agyat). A gépi jel-feldolgozás

²⁴ Ez az argumentum valójában egy változata Hubert Dreyfus sokak által osztott „klasszikus” AI-ellenes érvének, hogy tudniillik az emberéhez hasonló intelligenciához elkerülhetetlen, hogy az azt hordozó gépi entitásnak legyen(ek) saját „teste(i)”, amellyel a világhoz kötődik, és amelyre érvényesek annak törvényei, és legyen szocializációja, amellyel nemcsak képességei, hanem kontrollstruktúrái is felépülnek (Dreyfus, 1972, 1992).

²⁵ Emiatt – és a fogalom elterjedtsége miatt – „megtartjuk” az „intelligencia” kifejezést és az „intelligens” jelzőt, de avval a megjegyzéssel, hogy azt metonimikus értelemben használjuk, a megértést és a párbeszédet megkönnyítendő.

fejlődésének tehát többszörös tétje van: újabb és újabb területeken meghaladni a természetes intelligencia műveletvégző kapacitását, illetve annak alap-paramétereit (gyorsaságát, pontosságát, párhuzamos végezhetőségét, visszakereshetőségét stb.), és újabb és újabb területeken kiváltani a repetitív jellegű, alacsony érték-hozzáadású agymunkát.

A legendás és ikonikus Turing-teszt valójában mindkét célra alkalmatlan. Abból, hogy egy párbeszéd során embert látunk a velünk anonim módon kommunikáló gépben (még pontosabban: nem ismerjük fel a gépi intelligencia jelenlétét), nem következik sem az emberi intelligencia *utolérése*, sem *meghaladása*: legfeljebb feltételes, esetleges, ideiglenes és szituatív, tehát mindenképpen erősen korlátozott és szűk tartományra érvényes sikeres *szimulálása*. (És ez sem a gép autonóm teljesítménye, hanem a szimulációt irányító, programozó kutatóké).²⁶ S mivel a csevegés, a párbeszéd túlnyomórészt nagyon is magasrendű intellektuális művelet, az agymunka speciális fajtája, amelyen nincs mit „kiváltani”, egyre jobban látszik, hogy a gondolkodó gép metaforája²⁷ itt (és még sok más helyzetben is) értelmetlen, félrevezető és diskurzusromboló.²⁸

Bármilyen szuperintelligenciát tételezünk is fel, az mindvégig csak *az emberi értelem valamilyen (tehát nem minden) tartományban történő kiterjesztését* szolgálja. Nem általános/ge-

²⁶ Sokkal egyértelműbb mindez, ha Alex Tidemann doboló robotja, SHEILA van a fal mögött, aki neurális hálózata segítségével tanul meg egy dobszólót egyre jobban lejátszani, a dobos kézmozgását utánozva. Vajon ha a csevegés helyett az emberi és gépi dobolás közti különbség válik felismerhetetlenné, akkor SHEILA-t már intelligensnek nevezhetnénk? Aligha, mert nem tesz mást, mint pusztán reprodukálja az emberi dobjátékokat. Minderre lásd <https://www.youtube.com/watch?v=8SCWXbK42sc> (Letöltve: 2015. november 10.) Vagy itt van a tübingeni kutatók friss fejlesztése, a fényképeket adott festők stílusában átalakító algoritmus. Képfelismerő erő és a stílusjegyekkel kapcsolatos mintázatképzés arányos ötvözete kell ahhoz, hogy az eredeti művekre emlékeztető módon utánozható legyen az egyes festők jellegzetes, saját stílusa. A gépnek azonban ettől nem teremnek saját stílusjegyeik. http://index.hu/tech/2015/09/02/egy_ora_alatt_kesz_a_legujabb_van_gogh-re_mekmu/ (Letöltve: 2015. október 8.)

²⁷ Az ötvenes-hatvanas években, még kibernetikai fogatású narratívák részeként a „gondolkodó gép” nagyon hasznos segédfogalom volt, amellyel a szűk tudósközösség képes volt az érdeklődő közvélemény számára is érthető nyelvet alkotva népszerűsíteni a mindennapokba betörő számítógépek kultúráját és az abban rejlő lehetőségeket. Ma, egy sokkal nagyobb felbontású ismeretelméleti térben már nemcsak az érdeklődő közvéleményt téveszti meg, hanem magát a tudományos diskurzust is zavarja. A szuperintelligencia narratívája kerüli is a gondolkodó jelzést, mint a tüzet: meg kellene ugyanis magyaráznia, hogyan lehet valami úgy intelligens, hogy nem gondolkodik, hanem csak programot futtat le.

²⁸ Jól mutatja mindezt 2014 nyarának nagy híre, hogy tudnillik forradalmi áttöréssel sikerült megalakítani „a gondolkodó gépet”. A Eugene Goostmanre keresztelt program a zsűritagok több mint 33%-át megtevesztve hitette el magáról, hogy élő személy, egy 13 éves ukrán fiúcska. Csakhogy mindez részben a feltételek kijátszásán alapult (a fiatalság és az ukránság a hibaészlelési küszöböt szállította alacsonyabbra), a teljesítményt pedig nem gondolkodó gép, hanem egy chatbot, egy csevegésutatózó program adta le (amelyek közül a Cleverbot már 2011-ben is sokkal jobb eredményt produkált). Így is jellemző, hogy bombasztikus címek tudatták az „áttörést”, a „mérőldkövet a számítástechnika történetében” – lásd például az egyik legelső magyar híradást a PcFórumon (<http://pcforum.hu/hirek/16151/Forradalmi+attores+Megalkottak+az+elso+gondolkodo+gepet.html>), hogy aztán sokkal csendesebben adják át a helyet a kritikus és elemző, a kísérletet a megtevesztések közé száműző kommentároknak (például <http://bitport.hu/zavaros-siker-a-turing-teszten>).

nerális intelligencia tehát, hanem csak azokra az aspektusokra terjed ki, ahol a paraméter-növeléshez a számolás- vagy memóriateljesítmény növelésére van szükség. És nem a generális, hanem a szakosított irány tűnik adekvátnak: minél olcsóbb robotokkal minél jobban ellátni kicsi részfeladatokat (Brooks, 2003). Ez a „szűk mesterséges intelligencia” (*Artificial Narrow Intelligence, ANI*) programja, amelyet korábban a „gyenge AI” (*Weak AI*) fejezett ki.²⁹ Ám a korábbiaknak megfelelően még ezek is pseudo-intelligenciát jelentenek. Az emberi intelligencia ugyanis számtalan esetben nem komputációval dolgozik,³⁰ és evolúcióját már régóta eszközeibe helyezte át. Sem a szuperintelligencia, sem egy szűk területen elért előrelépés nem olvasható a gépi elem öntudatra ébredéseként, csak az egyik eszköztípus teljesítménynövekedéseként.

Semmivel nem vagyunk fogalmilag előbbre akkor sem, ha Luciano Floridi megoldását választva szimuláció helyett *emulációról* beszélünk (Floridi, 2014). A számítástechnikai környezetekben jól ismert emuláció-fogalommal ugyanis arra utalunk, hogy egy új (jellemzően: a korábbiaknál többet tudó) eszközkörnyezetben minőségromlás nélkül felismerhető, kezelhető, „átemelhető” egy korábbi eszközkörnyezetre szabott funkció. A lényeg tehát a kompatibilitás megteremtése, és ez kétségtelenül több, mint a pusztán másolás vagy utánozás. Csakhogy a mesterséges intelligencia nem az emberi értelem emulációja, nem az elmét „ülteti át” gépi környezetre. Pusztán kognitív részfunkciók válnak gépi úton is elvégezhetővé, s ha emulációt keresünk, arra csak ott bukkanhatunk, amikor a magasabb szintű teljesítményre képes gépi megoldás a korábbi, egyszerűbb megoldásoknál megszo- kott felületen is elvégezhetővé válik.³¹

Az alarmizmus reménytelenül félrevezető mivolta abból fakad, hogy tarthatatlan fogalmi kiindulópontonra építkezik. Nem veszi figyelembe, hogy ha számtalan művelet automatizálható is, nincs önmagában vett gépi intelligencia. A mesterséges intelligenciának (absztrakt gépi gondolkodásnak) ugyanis a mesterséges tudatosság (*Artificial Consciousness*) lenne az előfeltétele, amely viszont nem létezhet önazonosságra és önreflexióra képes ágens nélkül, amelynek identitástudata és állapot-tudatossága (*state awareness*) van. Az állapot-tudatosságnak párosulnia kell az állapotváltozás érzékelésére való képességgel, és

²⁹ Ide tartozónak tekintik a táblás játékok gépi rendszereit (Sakk, dáma, Scrabble, Backgammon, Othello), az IBM Watsonját, a Google fordítóprogramját, de akár egy spamfiltert is. Egy okostelefon olyan, mint egy „ANI-gyár”, számos „okos” alkalmazással (Urban, 2015).

³⁰ Hanem például hipotézisek felállításával és tesztelésével, amelyeket aztán a valóságban is kipróbál. Michael Littman szerint elvi lehetetlenség a valóságnak a valóságnál gyorsabb szimulálása: „*It is a logical impossibility that these computers would be able to accurately simulate reality faster than reality itself*” (Littman, 2015)

³¹ Vegyük a következő egyszerű okoskodást. A mesterséges intelligencia jelenlegi állapotában olyan, mintha azt kérdést tenné fel, hogyan tudnánk repülni, mint a madarak, ahelyett, hogy egyszerűen csak úgy fogalmazna: repülni akarunk. A formula, ahogyan a feladatot kijelöljük, hatalmas eltérésekhez vezethet a megoldásban. Még mindig nem építettünk a madarak repülését „lemásoló” légi alkalmazhatóságot (pedig sokáig kísérleteztünk csapkodó szárnyú gépekkel), ellenben sikerült olyan masinákat teremteni, amelyek gyorsabban repülnek bármely madárnál és eközben óriási terheket képesek szállítani. Az evolúció-formálta biológiai képesség és a technológia által alig néhány innovációs ciklus által megteremtett képesség közti különbségnek a mesterséges intelligencia-diskurzusokban való alkalmazhatóságáról szenvedélyes vita folyik, lásd például itt: <https://news.ycombinator.com/item?id=10165586>.

az állapotváltozást meghatározó tényezők azonosítására való képességgel. Mindehhez még a belső állapotra való szakadatlan referenciaképzés és az éppen aktuális külső állapot állandó összevetése is szükséges, egyedül ez lehet a jelentéstermelés alapja (Bodnar, 2015b).

Ezért nem jelenthetnek kiutat ebből a csapdából a „megváltásként” tálalt, a maga dimenziójában izgalmas eredményekkel kecsegtető *mélytanulás* (deep learning) koncepciókra épülő rendszerek, amelyek kognitív oldalról a konnekcionizmus, számítástudományi oldalról a neurális hálózatok korábbi paradigmáit próbálják ötvözni. A lényeg azonban ugyanaz, amin nem változtat az, hogy sokkal több adat alkotja a feldolgozás bázisát, hogy sokkal több réteg épül egymásra. Jól formalizálható, zárt jelrendszerekben, ahol a lehetséges válaszok az immár megfelelő számításteljesítmény birtokában algoritmizálhatóak, kiváló teljesítményre lehet képes egy deep learning technológia (nem véletlenül investálnak sokat ebbe az Internet nagyágyú). A képfelismerés és a beszéd felismerés bizonyos területein, radikálisan leszűkített funkcióterben kétségkívül sokkal fejlettebbek, mint elődeik, de ez csak annyit jelent – Bodnart parafrázálva –, hogy úgy eredményesebbek, hogy intelligensebb módon buták. S bár valamivel ennél is többet ígér az az irány, ahol a világot és annak összefüggéseit felfedező kisgyermek mintájára a szabályok és okoskodási rutinok nem előre programozottak, hanem a tanuló rendszer önmaga konstruálja meg azokat,³² az ágens tudatosság-állapotán mindez nem változtat semmit. Ahogy korábban a pusztán algoritmussal afféle parancsutásításos hadvezérként, most a tanuló algoritmussal van, türelmes tanítóként, nélkülözhetetlenül jelen az emberi értelem – hiába ígéri Pedro Domingos, hogy megalkotható a Mesteralgoritmus, amely mindent meg tud tanulni, annak köszönhetően, hogy öt különböző tanulási forma elemeit egyesíti (Domingos, 2015).³³

Az *aggregált gépi tanulásnak* (aggregated machine learning) nevezett irányzat sem más, mint a mélytanulás egyik lehetséges architektúrája. Ha például nem egy masinát tanítanak meg beszéd felismerésre, hanem sokat, mindegyik maga lép előre a megbízhatóan felismert és reprodukált nyelvi egységek számának gyarapításában, majd kicserélik, illetve egyesítik tudásukat, akkor sokkal rövidebb idő alatt növekszik meg a validált felismerő-készlet, mint ha egyetlen környezetben egyetlen rendszert fejlesztenének. Az aggregált tanulásnak is ugyanazok azonban a korlátai: felügyelt (supervised), hiszen a felismerés helyességét, az „elfogadást” csakis a természetes nyelvet jól ismerő emberek garantálhatják, másrészt a gyarapodó készlet mindig csak a megtörtént (aktualizált) és soha sem a lehetséges nyelvi aktusok és szókombinációk világára lesz érvényes (vagyis a hibaarány, még ha mindig csökken is, soha nem fogja elérni a 0%-ot). Funkcionálisan sokkal fontosabb a hiányzó néhány százalék beavadászásánál az, hogy ott, ahol a fizikai és/vagy szemantikai zaj ellenére hibamentes beszéd felismerésnek tétje van (például az ember-gép kommunikációban, ahol a gép a beszédutasítást végrehajtásba fordítja), ott ugyanaz az egy helyes felismerés ismét-

³² Ennek az iskolának az esélyeit leginkább Gary Marcus kutatásaival illusztrálják (Knight, 2015).

³³ S ugyanígy reménytelen az érző számítógép, az érzelmi intelligenciát formalizálni tudó mesterséges intelligencia megteremtésének programja. Egy emberi arc érzelmi állapothoz társításának képessége alakfelismerő mélytanulással elképesztő távolságra van attól, hogy a mesterséges ágens maga rendelkezzen a viselkedés-szabályozásban szerepet játszó érzelmi komplexummal. Ráadásul a humán ágens esetében az idegrendszer működésének két oldala, az érzelmi és értelmi szerves egységet alkot: egy evolúciós munkamegosztás részeként fejlődtek ki. Egyáltalán nem szükségszerű, sőt egyenesen indokolatlan, hogy a gépi intelligenciának is reflektálnia kelljen erre a bonyolult kettősségre.

lódhessen nagyon sok, egymást követő szituációban. Ez pedig valójában inkább a pedagógiában kialakult perszonalizált tanulási modellek átültetését jelenti a digitális környezetbe (egyetlen, eszközére folyamatosan ráhangoló emberre és egyetlen, rá fokozatosan „kalibrálódó” eszközre építve). Eközben a pusztá jel-aggregációban már „kollektív gépi tudatot” látni felelőtlen és félrevezető. Ám remekül teljesítenek együtt, ha a gépi elem az emberi tudás felhalmozásában segít. Amikor az egyes emberi tapasztalat formalizáltan leírható, és ennek révén valóban összeadódik és közkinccsé válik.³⁴

Hibriditás: a 'gépi' alapvető létállapota

A 'gondolkodó gép' fogalmával megragadott problématerben eltűnik az értelmezési tartományból az a tény, és nem elégszer hangsúlyozzuk, hogy nincs önmagában vett gépi intelligencia, az csak *hibrid (ember+gép) szerkezetben* tud megnyilvánulni, és az emberi mozzanat az elsődleges. Másrészt, mint azt a bevezetőben láttuk, ez nyit utat annak a pusztító és félrevezető diskurzusnak, amelyik a „*mikor éri utol és mikor múlja felül a gépi intelligencia az emberit*” ál-dilemmáját, vagy ennek még kakofóniába hajlóbb változatát, a „*mikor győzi le a gép az embert*” morális pánikba forduló kérdészt zenésíti meg.

Ez utóbbit bátran hívhattuk volna akár Deep Blue-effektusnak is, mert a korábbi, óvatos proféciai után a regnáló világbajnokot, Kaszparovot legyőző számítógép adta meg a bátorságot boldognak és boldogtalannak, hogy világgá kürtölje a lefegyverzően igaznak tűnő, valójában mégis veszélyesen félrevezető szenzációt: *a gép legyőzte az embert!* Egy ideje inkább Watson-effektusként érdemes már beszélni róla, amióta az IBM műveltségi játéokra kifejlesztett háromezer processzoros, tizenháromezer gigabájtos szörnyetegének sikere nyomán még a korábbinál is hangosabban rázzák a kereplőt a hagyományos és az online média leginkább szem előtt lévő felületein a vastag betűs, gondolatrest címek: *a gép ismét legyőzte az embert!*³⁵

³⁴ Ennek tipikus eseteit nem a Wiki-világban, hanem például a növény- és állathatározás, a gomba- és ásványfelismerés vagy az ornitológiai észlelések felhalmozása környékén kell keresnünk. A kollektív mozzanat itt azt is jelentheti egyúttal, hogy a besorolások/leírások validálását sem egyetlen szakértő végzi, hanem sokak egy irányba mutató megerősítése. Emiatt nagyon ígéretesnek tartjuk például azoknak a mobil alkalmazásoknak a fejlesztését, amelyek a személyi használatú eszközökön formalizálják az egyéni azonosítások felvitelét, majd az egymáshoz közel kerülő eszközökön szinkronizálják a gyarapodó adatbázist, fokozatosan építve ki egy kizárólag a felhasználók által alakított és egyre pontosabb „virtuális tudástárat” – amelyet in situ igénybe vehet minden felhasználó, ha számára ismeretlen azonosítási helyzetbe kerül. Az ismeretek így felfogott mesh hálózata nem a hozzáférést, hanem a tudás aggregálását segíti, s mivel ez tipikusan terepen történik, a mobil platform remekül igazodik hozzá (ilyen kutatások a Szegedi Tudományegyetemen folynak, Bilicki Vilmos vezetésével). A gépi környezetben aggregált emberi tanulás talán legfontosabb és legperspektivikusabb területét az *egyedi betegségleírásokból kinövő gyógyító praxis* körül kell keresni, és nem, ahogy Shawn DuBravac (2015) gondolja, az automata járművek által majdan generált majdani közlekedési helyzetek aggregált naplózásában. Itt sem gépi tanulás történne ugyanis, hanem az esemény-típusokból emberek formálnának új, még veszélymentesebb közlekedést lehetővé tevő algoritmusokat.

³⁵ És korántsem szükségtelen tudatosítani, hogy innen már olvasók milliói kerülnek egy lépésnyi közelségbe a falfehérré vált arccal, kezdődő pánikban elsuttogott baljós kérdéshez: *jaj, mikor fogja majd uralma alá is hajtani!?*

Vajon hányszor és milyen formában kell felhívni a figyelmet rá, hogy mindenki megértse: nem egy gép „győzte le az embert”,³⁶ hanem *egy programozó és mérnökcsapatból, valamint egy beszédfelismerő, jelfeldolgozó és beszédszintetizáló modulokkal rendelkező gépből álló hibrid rendszer*. Egy emberrel szemben tehát – mint egykor a sakkfejedelem túlóldalán is – egy páratlan művelet-végző sebességgel rendelkező masina és az azt megtervező, megépítő, programozó és strukturált tartalommal feltöltő szakértők tömege áll. Azért, mert a képernyőn a legendás Jeopardy-pult mögött „Watson” szimbolikus alakját látjuk két mesterszintű játékos között, el is felejtjük, hogy valójában emberek ármádiáit kell mögé képzelni, minden másodpercben? Ismételjük meg, hogy mi is történt valójában: emberi agyak egy tekintélyes, összekapcsolt csoportja alkotott és töltött fel tartalommal egy olyan szerkezetet, amelyiknek a szabályozott, kimenő jelei meghatározott keretrendszerben a sikeres emberi problémamegoldásra emlékeztető illúziót keltettek.³⁷

Sajnos még a legigényesebb jövőkutató irodalomban is visszaköszön ez a kettősség. Hiába vizionálja – helyesen – ember és gép új civilizációs minőséget eredményező fúziójának lehetőségét Ray Kurzweil, amikor szerinte az emberi agy „biokulturális” dimenzióban megragadható tudás- és képességpotenciálja ötvöződik az emberi tevékenység érdekében csatasorba állított okos eszközök nagyobb reprezentációs és műveletvégző kapacitásával, sebességével és információmegosztó képességével. Mindez a korábbi „szimbiózisoknál” kétségkívül szervezettebben forrasztja össze, utalja jobban egymásra, teszi nehezebben elkülöníthetővé a gépi és emberi komponenset (Kurzweil, 2013). Csakhogy téved, és ingoványos talajra kerül, amikor ehhez azt tartja szükségesnek, hogy „*a számítógépek elérjék a legmagasabb emberi intelligencia szintjét*” – a fenti szimbiózis minden további nélkül működhet enélkül is (sőt működhet nagyon hatékonyan).³⁸

Ez a tétel tehát jól működik, amikor a mesterséges intelligencia fejlesztése a kutatói kihívás, de ismeretelméleti szempontból zsákutca. A hibrid rendszereknek ugyanis éppen az a lényege, hogy *minden komponens a sajátlagosan rá jellemző képességmentésben adja le a teljesítményt, és ahol gyengébben teljesít, ott a munkamegosztás részeként átengedi a terepet*.³⁹ Az ár,

³⁶ Ráadásul miféle „győzelemről” is van szó? A hibrid rendszernek egy játékban aratott győzelméről, amely a játék természetes kimenete, akár humán-humán, akár gép-gép konstellációban. (Amikor az egyik sakkprogram legyőzi a másikat, akkor a gép legyőzi a gépet? Nem, az egyik programozó csapat bizonyul eredményesebbnek a másiknál). Csakhogy a semleges hangulatú, leíró jellegű „játékgyőzelem” a hétköznapi konyhanyelvben disztópikus jelentést ölt: ha a játékban a gép legyőzi az embert, akkor az már az előjele annak, hogy evolúciósan is felülmúlja, kiszorítja életteréből, aláveti, uralkodni fog rajta. És egyszer már csak ezen az absztrakciós szinten jelennek meg a gondolatok.

³⁷ Féltreértés ne essék: Watson produkciója fantasztikus és érdemdús mérnöki teljesítmény, a mesterséges intelligencia kutatásának egyidejűleg több, régóta mozdíthatatlannak tűnő pontján ígér előrelépést vagy áttörést. Elismerés és gratuláció illet mindenkit, aki a siker körül bábáskodott. Kritikánkat az interpretációval kapcsolatban fogalmaztuk meg, ami ráadásul hangsúlyozottan nem is a fejlesztőknek, hanem a bulvárosodó közbeszédnek köszönhetünk.

³⁸ S realiztikus jövőképe és tárgyilagos megközelítésmódja ellenére ezért csúszik át Kurzweil a science fiction irányába az elme fizikai testbe való áthelyezhetőségének tételével (amelyhez nemcsak az emberit utoléró gépi intelligenciára, hanem a testtel összekötött egyedi idegrendszerek működésének szimultán digitális replikációjára is szükség lenne).

³⁹ Az információs műveletvégzést támogató pre-digitális hibrid rendszerek is ezen az elven működtek. Az állati *jelző-riasztó rendszerek* (legyen az bányarígó vagy őrzőkutya) *kereséstámogató hibridek*

amelyet az így jelentkező koordinációs szükségletek ellátásáért fizetni kell időben, figyelemben és energiában, messze alatta marad a teljesítménynövekedésből származó előnyöknek. Ennek a folyamatnak (illetve sajátos, dinamikus egyensúlynak) a tervezésekor szokták az ember-számítógép viszonyra az *orkesztráció* (orchestration) kifejezést használni (Burgess, 2015, 341-345. o.), az egy időben leadott közös teljesítményre és a párhuzamos, illetve a sorba rendezett folyamatokra utalva. Három szakasza a *specifikáció*, a *kollaboráció* és a *koordináció*.

A gépi oldal fejlesztése tehát fontos terep, de csak egyike a négy alapvető kérdésnek: a *hibrid rendszer emberi oldalának fejlesztése* a második, az *emberi és a gépi együttműködésének kérdése (az interfész)* a harmadik, s az egész rendszerfejlesztésnek értelmet és célt adó *teleológia* a negyedik. Ez az a négy problémásík, amit minden pillanatban szem előtt kell tartanunk, s amely egyfajta osztályozási elvként működik: vajon az éppen mérlegre tett vagy vitatott elmélet a négy közül melyik szinten fogalmazódik meg, kellően komplex-e, értelmezhető-e egyidejűleg mindegyikre?

Foglaljuk össze tehát lista-szerűen a négy sík leglényegesebb jellegzetességeit, ahol lehet, szembesítve azt a meghaladni kívánt vagy vitatott nézőpontokkal.

1. A gépi oldal, a mesterséges intelligencia kutatói számára nem az emberi intelligencia elérése a valódi kihívás, hanem az, hogy a *mesterséges komponensre eső feladatok a legmagasabb szinten teljesüljenek, illetve szakadatlanul új feladatok váljanak teljesíthetővé*.⁴⁰
2. Az emberi oldal fejlesztése részben szociokulturális és oktatástechnológiai, részben pedagógiai, részben pszichológiai, részben pszichofarmakológiai kérdés,⁴¹ amelynek a következő időszakban sokkal nagyobb szerepet kell kapnia a diskurzusokban⁴² (s amelyre a kiborg-témakör rövid tárgyalása kapcsán majd vissza is térünk röviden).

(szarvasgombavadász disznó, nyomkövető vagy szaglókutya) vagy *üzenetküldő rendszerek* (postagalamb, futárkutya) például jól kiegészítették az érzékszervi és fizikai kapacitásában korlátozott emberi közösségeket, s a táplálás és – szükség esetén – betanítás költségei jóval alatta maradtak az együttműködésből származó hozadéknak. Jaron Lanier is észreveszi, hogy nemcsak „az emberi és a gépi”, hanem a „predigitális és a digitális” legjobb vonásainak szintézise az elfogadható irány. A legfrissebb kognitív tudományi eredmények is meggyőzően üzennek ennek a sajátos egyensúlynak a meglétéről. Sikerült például feltárni és bebizonyítani, hogy az adatok (és az általuk hordozott ismeretek) elmentése „külső hordozóra” felszabadítja a memóriánkat a következő hasznos információ befogadására – a „mentés” általi memória-tehermentesítés megnöveli a kognitív kapacitásunkat (Storm és Stone, 2015).

⁴⁰ Richard Boyd (2015) így fogalmaz: „*Hogyan érhetjük el a megfelelő egyensúlyt az emberi és automatizált között, hogy a kimenetet optimalizáljuk vele? Az emberi és a gépi intelligencia milyen kombinációja segíti legmegfelelőbben a legegyszerűbb problémák megoldását?*”

⁴¹ Itt részben a *nurture*, a tudatos nevelési környezet révén elérhető intelligenciafejlődés eszközvilágára gondoljunk, részben a memória, koncentráció stb. erősítését (ideiglenesen) stimuláló szerekre. Érdekes paradoxon, hogy a sikeres ember-gép kommunikációhoz olykor az emberi oldal „lebutítása” szükséges a gép szintjére (például primitív szintaxis a gép beszédértéshez), és nem a gépi oldal további „felokosítása”.

⁴² Az ember-centrikus megközelítés azonban nem jelenti azt, hogy akár az unalomig ismert luddita vagy az újkeletű „*humánsoviniszta*” megközelítésmód a legcsekélyebb mértékben is indokolhatóvá válna. Ez utóbbinak Jaron Lanier, a virtuális valóság atyja a legismertebb képviselője. Egy 2014-es díjátadási beszédében több egyéb tárgykör – például a nagy adat (Big Data) veszélyei – mellett egy olyan „új humanizmust” hirdetett meg, amely korábbi, a kibernetikai totalizmust (*cybernetic totalism*) elutasító deklarációi után azt a helyes állítást, hogy „az ember több a gépeknél és az algoritmusoknál”, összekapcsolja avval a helytelen következtetéssel, hogy emiatt a „mesterséges intelligencia

Emellett tömegesen „kiadhatóak” olyan feladatok az emberi elmének, amelyek megoldásában jobban teljesít, mint a gép – ez az alapja az *emberi számítástechnika* (human computation) irányzatának.⁴³

3. A két oldal közös felelőssége a megfelelő koordinációs minőség megteremtése a gépi és az emberi komponens között, hogy a hibrid rendszer a legnagyobb teljesítményt adhassa le.
4. S végül az egész rendszer működésének a *cél-vezérelt ko-evolúciós szemlélet* ad értelmet és megalapozást, nem pedig a mesterséges intelligencia-rendszerek öncélként felfogott „felgyorsítása” vagy az exponenciális ugrást lehetővé tevő „önfejlesztő üzemmódra” állítása. Mert e kettő mögé is fel kell tenni a kérdést, hogy mindez *mit szolgál a rendszer-egész szempontjából?*

HCI, humatika, kognitív infokommunikáció – visszaút a géptől az emberig

Első ránézésre a fenti felsorolás harmadik területével, a koordináció/interfész kérdéskörrel foglalkozik, ráadásul roppant kiterjedt módon és definíciószerűen, a HCI (Human-Com-

elutasítására” is szükség van. „*The new humanism is a belief in people, as before, but specifically in the form of a rejection of artificial intelligence*”. S még ha a szövegkörnyezetből ki is derül, hogy valójában a „mesterséges intelligenciához kapcsolt túlzott elvárásokkal” és az eluralkodó „fantáziavilággal” szemben foglal állást, és nem magát az algoritmusok világát tenné ad acta, evvel az egyetlen beszédével a zászlóvivőjévé és hivatkozási alapjává vált a kiber-utópizmust felváltó másirányú egyoldalúságnak. A kibernetikai totalizmus kritikájára lásd <http://hplusmagazine.com/2014/07/07/jaron-lanier-on-transhumanism/>, az új humanizmusról mondottakra, a teljes beszéd közzétételével: <http://www.friedenspreis-des-deutschen-buchhandels.de/819335/>. Az ember-közponitú szemléleti fordulat hasonlít ahhoz, amin (egy sokkal szűkebb mezsgyén) a humanitárius munkát végzők átmentek. Fel kellett ugyanis ismerniük, hogy hiába állítható elkötelezett tevékenységük szolgálatába számtalan nagyszerű technológia, a drónoktól a multiplayer játékokig. A digitális humanitarizmusnak tudnia kell, hogy az előretéktintő politika-alkotás és a résztvevők és vezetőik felkészültsége és felvilágosodottsága legalább annyira meghatározó (ha nem még fontosabb), mint a felhasznált technológia (Meier, 2015).

⁴³ Az immár szakosított intézettel (Human Computation Institute <http://humancomputation.org/>) és transzdiszciplináris folyóirattal (a Human Computation 2014 októbere óta jelenik meg <http://hcjournal.org/ojs/index.php?journal=jhc>) rendelkező irányzatot Luis von Ahn 2005-ös disszertációjára szokás visszavezetni. Ebben a guatemalai származású kutató és fejlesztő először fogalmazta meg alaptézisét a nagy tömegben önkéntes műveltségvezésre (crowdsourcing) fogott, emberi elmében rejlő komputációs erőről, amely megfelelő gépi környezetben olyan problémák megoldására lehet képes, amelyre külön-külön nem volnának elegek (nem utolsósorban a játékosítási és motivációs technikáknak köszönhetően). Azóta számtalan sikerprojektet sikerült tető alá hozni: az ExeWire-rel neurontérkép készülő (közel 200 ezer önkéntes segítségével), ígéretes az Alzheimer-kór kutatását segítő alkalmazás (WeCureAlz.com), ezekről részletesebben lásd Michelucci és Dickinson (2016). Tegyük hozzá a teljesség kedvéért, hogy a fogalom születésének évében, 2005-ben jelent meg David Alan Grier könyve a „humán komputerekről”: azokról az emberekről (főleg nőkről), akiket a huszadik század fordulóján iparszerűen alkalmaztak arra, hogy nagy tömegben végezzenek el számológépműveleteket (Grier, 2005). A számítógép megszületése ezeket a humán komputereket váltotta ki. De ma ott, ahol a gépek tudása véget ér, vagy az ember hatékonyabb lehet (például: mintázat-felismerés és osztályozás, kreatív absztrakció) egy magasabb szinten (és az egykori irodamérethez képest a feladatok sok kis részre bontásának, a microtaskingnak köszönhetően sokkal nagyobb tömegek bevonását lehetővé tevő módon) térnek vissza.

puter Interaction), az ember-számítógép kapcsolat (ESZK) tudománya.⁴⁴ Csakhogy a HCI elsősorban nem a ko-evolúcióra és annak teleológiájára koncentrál, magas absztrakciós szinten, hanem az adott funkcióra létrehozott gépi környezetek alacsony absztrakciós szintű, felhasználó-orientált hatékonysági és ergonómiai fejlesztésére.⁴⁵ Jól tükrözi ezt a HCI saját történeti identitása is: a fogalom első, szórványos előfordulását 1975-re teszik, a diskurzusnyitó monográfia (Card et al., 1983) megjelenését alig több mint 30 évvel ezelőttre. Eközben a ko-evolúciós gondolat évtizedekkel korábban, a hatvanas évek elején, nagyon is jövőtudatos formában vetődött fel: a Világháló egyik korai építője, J.C.R. Licklider vezette be a „szimbiotikus rendszer” kifejezést az ember-számítógép kapcsolatra (Licklider, 1960, 1965).⁴⁶ A számítógép-komponens feladata nála nem az emberi intelligencia utolérése vagy meghaladása, hanem *augmentációja* (feljavítása), és a szimbiotikus ember-gép rendszer közös célfüggvénye a (közös) intelligencia *amplifikációja*, felerősítése (intelligence amplification, IA). („Az ember jelöli ki a célokat, szabja meg a feltételeket, és végzi az értékelést. A számítógépek végzik a rutinizálható munkát”). Lickliderrel közel egyidőben szintén az augmentáció fogalmában ragadta meg a kihívást Doug Engelbart: kifejezetten a komplex szituációk megértésének és megoldásának eszközeként (Engelbart, 1962). Sok évtizeddel később John Thackara egyenesen a HHI (*Human-Human Interaction*) kifejezésre való váltást javasolta (Thackara, 2005) annak az érzékeltetésére, hogy *még ott is a humán oldal a lényeges, ahol dramaturgiailag fontos szerepben van jelen a gépi komponens* – mint például a hálózatokban. Ahogy Brynjolfsson és McAfee (2014) fogalmaznak: a brutális processzor-erőnek az emberi leleményességgel kell párosulnia.

Engelbart jellegzetes tipológiája, amely a humán oldal „feljavításának” lehetséges formáit vette számba, mai napig érvényesen jelöli ki az irányokat:

- gyorsabb megértés;
- jobb megértés;
- a megértés megfelelő szintjének elérése korábban túl komplexnek tűnő probléma esetén;
- gyorsabb megoldás;
- megoldás találása olyan problémákra, amelyek korábban megoldhatatlannak tűntek.

Jól látszik, hogy *a megértés a humán komponens célfüggvénye, a megoldás a hibrid rendszeré*. Mindez semmilyen formában nem implikálja a gépi oldal egyoldalú kiemelését. Sokkal inkább következik belőle az, hogy ha a hibrid rendszer gépi komponense „előreszalad”, nagyot ugrik teljesítményben, akkor az emberi oldal felzárkóztatása, „mérétezése”, a megfelelő kapcsolat kialakítása válik meghatározóvá a harmonikus fejlesztés érdekében. Ve-

⁴⁴ Használják még az ESZI (Ember-Számítógép Interakció) és – főleg az ergonómusok – az EGK (Ember-Gép Kapcsolat) rövidítést.

⁴⁵ Minderre egyre népszerűbb az interakció-tervezés (*Interaction Design*) kifejezés is, amelynek célja a zavaró mozzanatok kiküszöbölése és a pozitív élmény elősegítése és fokozása (Lásd például Preece és társai (2015) tíz év alatt immár negyedik kiadásban megjelenő tankönyvét).

⁴⁶ Mivel Licklider és munkássága kevésbé volt látható, ezért a gondolatot sokáig az üzleti számítógép-fejlesztés egyik pionírjához, John Dieboldhoz kötötték (Diebold, 1969), aki azt javasolta, hogy minden lényeges kérdést az „ember és számítógép” egységeként közelítsünk meg.

gyük észre például, hogy gyorsabb és jobb megértés érdekében számos technikát állíthatunk csatasorba, amelyek *nélkülözik a gépi elemet*.

Az újabb generációs infografikai szcéna például pontosan arról szól, hogy a numerikus nehézbombázással átláthatatlanná tett vagy túl sok változó egyidejű figyelembe vételét igénylő tudás-mikroverzumok művészi és professzionális megjelenítésével gyors, holisztikus, elmélyült és élményszerű megértést biztosíthatunk. Az augmentáció forrása lehet egy vagy több másik elme nagyszerű teljesítménye is, amely sok esetben gépi és automatizált adatgyűjtések és feldolgozások kimeneteit fordítja át használható tudássá. Az óriási adattömeg szemantikus operátorok segítségével elvégzett tartalomelemzésén alapuló előrejelzéseknél sikeresebbnek bizonyulhat a „tömegek bölcsességének” igénybe vétele.⁴⁷ Az is igaz ugyanakkor, hogy a komplex ismeretrendszerek vizualizálásának technológiája vagy az online önkénteseknek teremtett platform már ismét a gépi oldal felé mutat: a vonatkozó számítógépes képszerkesztő és megjelenítő alkalmazások, animációs lehetőségek és workflow eszközök amplifikálják az új értéket létrehozó emberi tevékenységet.

Ha a HCI funkcionális történetére vetünk pillantást, azt látjuk, hogy létrejöttekor és első szakaszában a meghatározó terület a *biztonságkritikus rendszerek* világa (erőművek, repülés) volt. Ebben a szakaszban a kisebb és nagyobb rendszerzavarok elemzése nyomán rendre kiderült, hogy a kiküszöbölendő hibák nagyrészt *az emberi komponens gyengeségére* (elfáradás, figyelemhiány, rossz döntések sorozata, előírások és karbantartási feladatok figyelmen kívül hagyása, frissítés elmaradása stb.) vezethetőek vissza, részben az evvel nem számoló *tervezés hibái*: a gépi oldal jellemzően megbízhatóan teljesített.

A HCI történetének második szakaszában a tömegfelhasználás került a középpontba: diszciplináris oldalról az ergonómia és a design (s annak révén az artisztikum, a művészi kreativitás és invenció). „Segédtudományi” oldalról a viselkedéskutatás, fejlesztési oldalról az új beviteli és visszacsatolási eszközök. Praktikusan: a felhasználók szempontjainak figyelembe vétele már a fejlesztési szakaszban – olyan módszerek felhasználásával, mint például a pszichológiából importált Q-metodológia.⁴⁸

Amellett érvelek, hogy elérkezett az idő a HCI harmadik szakaszára, amelyet legszívesebben *koordinált jövőtudatos szakasz*nak neveznék. Az elnevezés elsősorban azt tükrözi, hogy a második szakaszban (és evvel párhuzamosan: az Internet-korszak utóbbi húsz évében) az eszközfejlesztői, -gyártói logika határozta meg a gépi oldal fejlődését, és az akadémiai jellegű HCI-kutatások is jórészt elszigetelt, egymással versenyben álló hardver- és szoftverfejlesztő üzleti vállalkozások innovációs segédcsapataként üzemeltek. Mostanra alakultak ki az előfeltételei annak, hogy megkezdődjön egy egészen más elvekre és kiindulópontokra épülő kutatási szakasz, amelyben a tervezői gondolkodás (design thinking)

⁴⁷ A négyéves kutatásra épülő Good Judgment Project több ezernyi önkéntese a geopolitikai előrejelzésben meglepő pontosságot ért el, és lekörözte minden mesterséges intelligencia-alapú megoldás eredményességét – kiemelve annak fontosságát is, hogy ahol működik, lehet az első gondolatunk a humán intelligencia a gépi helyett (<http://www.goodjudgmentproject.com/index.html>).

⁴⁸ A pszichológus William Stephenson (1902-1989) által kifejlesztett módszertan a „szubjektivitás” mérésére szolgál, ennek HCI-re alkalmazott változata, a HCI-Q az iteratív design-ciklusokban használható, amikor a felhasználók és mások szempontjai felől veszik szemügyre a fejlesztett eszköz személyes szignifikanciáit (O’Leary, 2013). A kérdéskör 15 oldalas bibliográfiáját Charles H. Davisnek köszönhetjük (Bibliography of Qmethodology in audience research). <http://www.ryerson.ca/~c5davis/Q-studies-of-audiences.pdf>

a jövő-írásstudással (futures literacy) találkozva az alapoktól gondolja újra, s kezdje meg újraépíteni, majd orkesztrálni azt, ahogyan az emberi elme és az azt amplifikáló gépi intelligencia összekapcsolódik.

Hans-Dietrich Kreft az ezredforduló környékén „humatikának” nevezte el azt az irányzatot, amely még a szenzorok, szoftverek és autonóm rendszerek gépi világát is kizárólag az emberi környezetbe ágyazott, illetve az emberi életet gazdagító mivolta révén tartja tárgyalandónak.⁴⁹ S mivel az ember bármilyen technológiai rendszer lényegi középpontja, mindvégig a tudás interoperábilis fizikai jellegzetességeinek megértése, leírása – praktikusán: cserélhetőségének és mérhetőségének megteremtése – lesz a lényeges mozzanat (Kreft, 2003).

2015-ben, legnagyobb örömünkre, számos olyan könyv is napvilágot látott, amely Kreft kiindulópontjait viszi tovább, immár a legfrissebb technológiai környezetre alkalmazva. David A. Mindell a jelen technológiai rendszereinek középpontjában álló gazdag emberi jelenlét (*rich human presence*) tézisét helyezi szembe a mesterséges intelligenciával párosult robotok autonómiájával kapcsolatos tévképzetekkel (Mindell, 2015).⁵⁰ John Markoff pedig egyenesen odáig megy, hogy emberi és gépi koevolúciójában nem a mesterséges intelligenciára (AI), hanem a humán intelligencia feljavítására (intelligence augmentation, IA) helyezi a nagyobb hangsúlyt: a gépi elemnek ez utóbbit kell szolgálnia, hogy önmagunk, emberi mivoltunk újfajta „design”-jával készüljünk a jövőre (Markoff, 2015). És ez nemcsak biológiai, hanem társadalmi létünkre is igaz: Kentaro Toyama központi tézise, hogy a technológiai fejlődés mit sem ér, ha annak vívmányai nem vonhatóak be az akut társadalmi gondok megoldásába. Ezt a változást pedig az emberi bölcsesség és tudás irányíthatja csak, és ennek a kapacitásnak a megnövelése, és nem a helyettesítése a mesterséges intelligencia köré épülő technológiánk küldetése (Toyama, 2015). Kulcskérdéssé az emberi elme fejlesztése, a tudós- és tudásközösségek együttes szellemi teljesítményének (Chris Anderson szavaival: raj-intelligenciájának) fokozása válik, és mindebben felértékelődik a Turing és Neumann-típusú géniuszok szerepének megtalálása is (Hsu, 2015; Colvin, 2015).⁵¹

És ez az a pont, ahonnan a magát *kognitív infokommunikációnak* (Coginocom) nevező irányzat létjogosultsága is leginkább megérthető és megérthető.⁵² A 2006 óta, javarészt

⁴⁹ Cége, a Humatics Corporation (<http://site.humatics.com/>) ezekre az elvekre építi profilját, amellyel piaci szereplők számára nyújt tudás-szolgáltatásokat.

⁵⁰ Mindell egyúttal azt a rendkívül fontos szempontot hangsúlyozza, hogy intelligens és önvezérlő mechanikai eszközeink az emberi tevékenység határait terjesztik ki extrém környezetekben. Basulto (2015) pedig méltán figyelmeztet arra, hogy *a robotok evolúciójának* (robotic evolution) is éppen ez ad értelmet és célt: életterek benépesítése, a megismerés határainak kiterjesztése, és az emberi cselekvés korlátainak megszüntetése. Frank Tipler pedig bátran ki is mondja: a mesterséges intelligencia küldetése nem más, mint a hozzájárulás az emberiség megmentéséhez és az új kolonizációjához (Brockman, 2015). Ehhez képest az *evolúciós robotika* (evolutionary robotics), amely darwini trükköket alkalmaz a gépi intelligencia fejlesztéséhez, „csak” egyike a versengő fejlesztési paradigmáknak.

⁵¹ Brian Eno egyenesen azt ajánlja nekünk, hogy képzeljük el az emberi társadalmat, mint a legerősebb szuperszámítógépet, és magát a globális civilizációt, amely mesterséges intelligenciaként ölel körül minket (Brockman, 2015). Más kérdés, hogy evvel a huszadik század harmincas éveitől íródó „világagy-diskurzushoz” kanyarodunk vissza.

⁵² Érdekességképpen említsük meg, hogy a nagy szóalkotó, Stanislaw Lem lexikonában a fejlett mesterséges intelligenciára használt technoevolúció (technoevolution) és intelletronika (intellectronics) mellett született kifejezés a kognitív teljesítményt növelő technológiákra is: a cerebromatika (cerebromatics).

magyar kutatók kezdeményezésére elindult interdiszciplináris útkeresés⁵³ abból indul ki, hogy a komplex technológiai- és médiatérben minden kommunikációs aktusok révén történik, emiatt a kognitív tudományban felhalmozott tudást kell párosítani a mérnöki alkalmazásokkal (Baranyi, 2012). Másképpen: a jelenlegi eszközvilágot a lehetőségek maximumáig kell a kognitív folyamatokhoz illeszteni, hogy a kommunikáció hatékonyságának növelésével javuljon a rendszerteljesítmény is. Evvel a normatív, metaelméleti szint újra összetalálkozik a praktikus fejlesztési filozófiákkal: ideje hát kibővíteni a valójában leegyszerűsített ember+gép modellt, mert a szemléleti változás szükségességének legmélyebben fekvő okai csak a teljes kapcsolatrendszer ismeretében tárhatóak fel.

Ember és számítógép – út a kiterjesztett modell felé

Eddig ember-gép szimbiózisról beszéltünk, de az önmagában vett „hibrid” rendszerről, absztrakt módon beszélni értelmetlen: hiszen az ember-gép együttműködés okát és értelmét kizárólag konkrét helyzetekben, annak megnyilvánulásakor tudjuk értelmezni. Ember és mesterséges intelligencia párosáról tehát ideiglenesen, egy bonyolultabb építmény „alsó szintjeként” beszélhetünk, amely kizárólag a ráépülő elemekkel együtt lehet érvényes.⁵⁴

(EMBER + AI)

A gépi intelligencia csatasorba állítása kivétel nélkül olyan élethelyzetekre reflektál, ahol az augmentáció igénye felmerül.⁵⁵ Minden monoton számolásművelet értelme az eredmény lehetőleg azonnali „lefordítása” valamilyen valóságos problémára (nem véletlen, hogy a számítógépek tömeges felhasználása a vállalatok bérszámfejtési osztályain és a pénzügyintézetekben indult meg). Az „emberi komponens” kizárólag cselekvő emberként van jelen a rendszerben,⁵⁶ emiatt a cselekvésekkel (illetve azok összességével, a viselkedéssel) ki kell bővíteni a modellt. Ha a mesterséges intelligencia hozzájárulása a cselekvés hatékonyságát fokozza, akkor *a különböző hibrid rendszerek különböző cselekvéstípusokhoz rendelődnek hozzá*. Másképpen: a hibrid rendszerek egyúttal funkcionális rendszerek, ahol a közös működés értelmét és „behuzalozását” adott funkciókból levezethető feladatok és az ezekre reflektáló cselekvések vezérlik.

⁵³ 2015-ben már a 6. nemzetközi konferenciát rendezték a tárgyban, egyre növekvő érdeklődés mellett. Az érintett tárgyköröket, a fő kategóriákat jól tükrözi a végleges program. (<http://coginfocom.hu/conference/CogInfoCom15/>) Az irányzat intézményesedését a felsőoktatás képzéskínálatában való erőteljes megjelenés is jelzi (szakirányok, doktori témakiírások formájában).

⁵⁴ E modell egy korai, fejletlenebb változat (Z. Karvalics – Juhász, 2008) alapos kibővítése.

⁵⁵ Hangsúlyozzuk ugyanakkor, hogy ennek nagyon fontos járulékos következménye az az állítás, hogy sok élethelyzetben nincs szükség augmentációra. Így minden univerzális augmentációs próbálkozás eleve feleslegesen túl nagyra tágtított teret kíván betölteni.

⁵⁶ Vannak első pillantásra extrémnek tűnő kivételek, például a nyugalomban lévő vagy alvó ember életműködésének paramétereit monitorozó szenzorrendszerek, amelyek automatikusan továbbítják mobil adatátviteli moduljuk segítségével a kimenő jeleket az azokat értelmező feldolgozó központba. Csakhogy ezek a megoldások valójában az orvosi cselekvés hatókörét kiterjesztő megoldások, a zavarok korai észlelésének funkciójával, amely a működés értelmét adja.

(EMBER + AI) ← FUNKCIÓ

Míndez tárgyszinten is azt jelenti, hogy a funkció határozza meg a hibrid rendszer gépi komponensének mineműségét: teljesítményét, hangolását, kimenetének szabályozását, az emberi komponenssel való kapcsolódásának mikéntjét és felületét, az interfészt. Emiatt az összefüggést akár így is felírhatnánk:

EMBER + (AI ← FUNKCIÓ)

A hibriditásnak értelmet adó funkció tehát a cselekvés jellege felől határozza meg a támogatási szükségletet. Emiatt aztán azonnal két ágra szakad az informatikai univerzum is: az információs viselkedést támogató rendszerekre⁵⁷ és a fizikai jellegű cselekvésekhez (mozgás, objektumra irányuló mechanikai műveletek) kapcsolódó rendszerekre.⁵⁸ Ez utóbbi esetben az ember-információs gép hibrid kiegészül egy mechanikus komponenssel is,⁵⁹ amely az anyagtudomány és robotika világával eredményez újabb és újabb fúziókat (a beágyazott rendszerek újabb generációjának megteremtésével).

Csakhogy a modell még mindig csak részleges. Az egész hibriditás végső értelmét azok a (legtágabb értelemben vett) környezeti beágyazottságok adják, amelyekből a funkció fakad: emiatt a platformok esetleges közössége vagy azonossága ellenére határozottan elkülönülnek egymástól a *társadalmi* (közösségi és épített), a *természeti* (geológiai, biológiai és kozmikus) és a *szimbolikus* környezet kihívásaira reflektáló megoldások. Olyannyira, hogy tulajdonképpen *három alternatív mesterséges intelligencia-fejlesztési irányról, alap-paradigmáról* kellene beszélnünk, mert a különbségeiket elfedi a „homogén” AI képe. A különbségekre érzékeny megközelítmődnak az ad különös jelentőséget, hogy a „Minden dolgok Internetje” (Internet of Everything) világában ezek a különböző törvényszerűségekkel jellemezhető AI-kisvilágok ismét összekapcsolódnak, hosszú operációs láncok, „kaszádok” részeiként.

Egyének és csoportjaik e különböző környezet-dimenziókhöz való egyidejű viszonyukból származtatják a támogatási igényt, életműködésük optimalizálását illetve javítását eredményező változók meghatározásával.

EMBER + (AI ← FUNKCIÓ) ← KÖRNYEZET →

Azt is mondhatnánk, hogy innen érthető meg a „gépi oldal” *teleológiája*: az életre hívásának, fejlesztésének értelmet adó – és ezért felépítését, működését, fejlesztését alap-

⁵⁷ És itt is az információs viselkedés három különböző fajtájának szükségleteihez igazodó módon. (Belépő oldal: információszerzés – feldolgozás – kilépő oldal (tárgyasítás).

⁵⁸ Természetesen az információs viselkedésnek is mindig van fizikai komponense, és a fizikai cselekvésnek is információs komponense, ezekkel azonban „egyszerűsíthető” az alapmodell. A két rendszer határára állnak az úgynevezett beszédaktusok, ahol a kimondott szó „értéke” és funkciója megegyezik a fizikai cselekvésekével (simogatás, bántalmazás stb.).

⁵⁹ Ezek története a szívritmus-szabályozókkal, a pacemakerekkel indul, és a gyógyászatban, katonaságánál és nehéz raktározási feladatoknál alkalmazható exoskeletonoknál jár, természetesen ide sorolandó a „dolgok Internetje” és minden, ami a közlekedés, a szállítás és a gyártás informatizálása révén koordinálhatóvá és vezérelhetővé válik.

vetően meghatározó, cél-természetű – kiindulópontok az ember-környezet kapcsolatból származnak. Ugyanakkor a cselekvési ciklusok is minden esetben a környezetre hatnak vissza, így aztán a „kibővített hibridet” ekképpen tudjuk megjeleníteni:

(EMBER ↔ KÖRNYEZET) + (AI ↔ FUNKCIÓ)

Az innen kinyerhető összefüggések csak látszólag banálisak és csak részben fedik át egymást a „*technológia társadalmi konstrukciójának*” (*Social Construction of Technology*, SCOT) vagy még inkább a *cselekvőhálózat-elmélet* (*Actor–Network Theory*, ANT) téziseivel illetve irodalmi hagyományával – ezekre együtt lásd (Király, 2005).

Úgy is mondhatnánk, hogy az ANT a fenti „képlethez” egy új dimenziót rendel, hiszen az ember+gép hibrid helyébe egy sokkal bonyolultabb, cselekvő entitást állít, amelynek egyaránt részei különböző emberek, technikai objektumok, más „materális” elemek és még jelentések is (szemiotikai objektumok). A hálózat (amelynek node-jai, csomópontjai is hálózati természetűek, s más hálózatok által egyszerűsítettnek, és más hálózatokat egyszerűsítene) az emberi összetevőt is a hálózatok által formált entitásnak láttatja: így nemcsak a valóságba beavatkozó (és abba rendet vivő) cselekvéseket, hanem még a jelentéseket is a hálózat generálja. Emiatt a hálózat minden entitását (legyen az emberi vagy nem emberi) ugyanazon a módon lehet megragadni és leírni (az ANT szótárában ezt fejezi ki az általánosított szimmetria (*generalized symmetry*) elve.

S ha a cselekvőhálózat maga egy elképesztően bonyolult, finom szövésű rendszer, akkor elképzelhetjük, hogy változása, módosítása, netán átalakítása mennyire összetett egyensúlyi térben értelmezhető. Ahogy Callon (2005,101) magyarázza: „*minden módosítás a cselekvő hálózat elemein és kapcsolatain kívül érinti az egyes elemek által egyszerűsített hálózatokat is... az átalakítás így a cselekvő hálózatot alkotó különböző elemek ellenállásának vizsgálatán múlik*”. A társadalomtudomány hibás előfeltételezésekkel igyekszik megérteni és előre jelezni változásokat: „*szükségképpen hipotetikus és spekulatív fog maradni, mivel a társadalmi realitás egyszerűsítése közben a vizsgált asszociációk közül elhagyja mindazokat az entitásokat..., amelyek képesek megmagyarázni a társadalomnak és termékeinek koevolúcióját*”. De mindez fordítva is igaz: ha egy elmélet kizárólag a technológiai komponens változásaira érzékeny (mint az erős mesterséges intelligencia és a szuperintelligencia tézise), szükségképpen hipotetikus és spekulatív fog maradni, mivel a technológiai realitás egyszerűsítése közben elhagyja azoknak a cselekvő hálózatoknak a más entitásait, amelyek csak együttesen képesek válaszolni a feltett kérdésekre.

Még jobban látszik mindez, ha felismerjük, hogy az ANT figyelemre méltó diskurzusában a mesterséges intelligencia valójában legalább négyféle formában entitásképző.

1. *Az emberek helyére szituatívvan ember-gép hibrideket kell állítanunk.* Ugyanezek az emberek ettől még a gépi elem nélkül is, és ugyanazok a gépek az emberi elem nélkül is hálózati pontok, és más helyzetekben más elemekkel alkotnak funkcionális rendszereket.⁶⁰ Vint Cerf és David Nordfors egyenesen erre az entításra szabták a víziójukat.⁶¹ Olyan

⁶⁰ A funkcionális rendszerek Anohin-féle elméletét termékeny módon lehet a hálózatokkal (jelesül: a minden dolgok Internetjével) összekapcsolva tárgyalni (Z. Karvalics, 2015).

⁶¹ <http://i4j.info/2014/07/disrupting-unemployment/1502/> (Letöltve: 2015. szeptember 1.)

- mesterséges intelligencia-rendszerek fejlesztését sürgetik, amelyek személyre szabottan és külön-külön segítik a bolygó minden egyes lakóját abban, hogy az általa kedvvel végzett tevékenységekhez kapjon külső támogatást (s ha ez a kapacitástömeg kiváltaná és felülírná a gyűlölt, kényszerből végzett munkát, akkor avval mind a gazdasági növekedés útakadályai, mind a foglalkoztatás krízise kezelhetőek volnának).
2. *A mesterséges intelligencia-rendszereket fejlesztő tudósközösségek tagjai és az általuk orkesztrált gépi komponensek együttesen hálózati természetűek*, de hálózati entitásként *csomópontszerűen* kapcsolódnak más hálózatokhoz (például kutatói közösségekhez általában, vagy a konkrét alkalmazási-felhasználási területek emberi és nem-emberi komponenseihez, vagy olyan gazdasági szereplőkhöz, akik a fejlesztés vagy a piacra vitel kontextusában finanszírozók vagy potenciális vásárlók).
 3. *Bármely AI-objektum (szoftver- és hardver-komponenseivel) önmagában is cselekvő-hálózati entitás*, és önmagában is lehet hálózati természete (raj-intelligenciája például). A legbonyolultabb, akár több hálózati „réteget” magába foglaló AI-szuperrendszer is azonnal pontszerűvé válik azonban, amint műveletet végez, hiszen avval a cselekvő-hálózat részévé válik.
 4. S végül *az AI-diskurzus maga is része a cselekvő-hálózatnak*: azok a fogalmak, terminusok, jelentések, képzetek, referenciák, asszociációk (nevezzük az AI szemiokulturális dimenziójának) alkotnak önmagukban is hálózatot, amelyekkel az AI kérdései felé fordulunk. Amelyek alapján az emberi komponensek általános és speciális viszonyokat alakítanak ki, döntéseket hoznak, és cselekvéseiket hozzáigazítják ezekhez a mentális tartalmakhoz. Ez az a mozzanat, amely ismételten felértékeli az alarmista nézőpontok kritikáját – annak tudatában, hogy a *nézőpont* egy cselekvő-hálózatban egyúttal *csomópont*, amely ugyanúgy formálhat kimenetet, mint bármely materiális komponens.

És erről a teraszról látszik csak igazán, mennyire leszűkül az erős mesterséges intelligencia és a szuperintelligencia teleológiája a harmadik típus, az AI „önmagában vett”, cselekvő-hálózatból kiemelt, steril értelmezésére. Az „emberit elérő és meghaladó” intelligencia-entitás létrehozása kétségtelenül öncél: az ezt elkerülhetetlennek láttató előrejelzések pusztán technológiai dinamikák, felgyorsulások és trajekciók mesterségesen létrehozott, hálózati beágyazásukból kiszakított vákuum-világában értelmezhetőek. Az, hogy a diskurzusban ez gyakran összekapcsolódik a várható „jótéteményekkel” (gyógyítás, tudományos problémák megoldása) nem feledteti, hogy e kapcsolódásnak szervesnek, funkcionálisnak és nem hierarchikusnak illetve kauzálisnak kell lennie.⁶² Ahogy Bruno Latour fogalmaz: „*a kutatási programoknak oly módon kell asszociálniuk egymással elméleteket,*

⁶² Egy egyszerű példával illusztrálva: az öncélú mesterséges intelligencia-fejlesztés irányt vehet például egy betegség-típusra fogékonyra tevő, genetikai kapcsolat-mintázatok feltérképezéséből származó, „nagy adatokból” új algoritmusok segítségével kinyert heurisztikák megtalálására – miközben ez csak az egyik lehetséges út a betegség leküzdésére. Mi van, ha sikerül a rizikófaktorok visszaszorításával eliminálni a betegség kialakulásának veszélyét? És mi van, ha egy etnomedicina alkalmazásával sikerül megoldást találni a már kialakult betegség villámgyors kezelésére? De gondolhatunk akár az Internetre is: korai fejlesztését a hatvanas-nyolcvanas években *akadémiai teleológia* vezette. (A létező, de nem egyedül meghatározó *katonai-hidegháborús* „szál” alapvető szerepe csak széles körben elhíresült városi legenda. Ennek forrása az volt, hogy a hálózat pionírjai a sok lehetséges útvonalat

fogalmakat, osztályozásokat, nem embereket,⁶³ megfigyelési eszközöket és technikákat, intézményeket, világnézeteket vagy akár politikai cselekvőket és ideológiákat, hogy azok ne mondjanak ellent egymásnak” (Berger, 2008, 83. o.).

Az ANT azonban csak egyike az elemzési-leírási keretrendszereknek. Önmagában nem ad sem teljes magyarázatot, s nem vezethetőek le belőle automatikusan a jövő alakításával kapcsolatos normatív elvek sem. Emiatt az „egydimenzióssá” silányított, a szingularitást axiómává emelő technológiai diskurzus mellé egy átfogóbb, a civilizációs teleológia birtokában „újratervezett” diskurzusra van szükség – és úgy tűnik, alarmisták ide, morális pánik oda, a folyamatok, az emberek és az erőforrások ennek az irányynak megfelelően kezdenek összeszerveződni.

A Future of Life Institute 2015. január 11-i nyílt levele⁶⁴ manifesztum-szerűen foglal állást az AI-nak a társadalomra gyakorolt pozitív hatása és gazdasági értékei miatt szükség-szerű, koordinált fejlesztésének fontossága mellett. A nyílt levél melléklete (*Research priorities for robust and beneficial artificial intelligence*)⁶⁵ strukturált módon foglalja össze, milyen sarkalatos kutatási témák köré rendeződhetnek a jövő tudományos és technológiai erőfeszítései, a tárgykör egyfajta problématerképét (és az ahhoz igazodó, közel száz tételes szakirodalomlistát) kínálva (az írás magyar változata tematikus számunk végén olvasható – a szerk.). S biztató, hogy a kutatók, fejlesztők, üzletemberek, újságírók (a csatlakozásra továbbra is nyitott nyílt levél aláírói) között a legnagyobb „vészmadarakat” is ott találjuk. Az is előremutató, hogy a szerzők tisztában vannak vele: önmagában a mesterséges intelligencia kutatása nehezen lép már előre a gépi tanulás, a statisztika, az irányításmélet vagy a neurológia tudományaival való szövetségek nélkül. Más kérdés, hogy amikor a fő alkalmazási területeket tekintik át, akkor – az önvezető járműveken kívül – rendre csakis régi, jól ismert darabokat sorolják fel: beszédfelismerés, gépi fordítás, kép-osztályozás, kérdező-válaszoló rendszerek, lépkedő robotok stb. Ezek kétségkívül erősen foglalkoztatják azokat, akik a mesterséges intelligencia fejlesztésének tudományos és üzleti hátországában dolgoznak.

A mindennapok hullámverésében azonban más válik fontossá. Amiként a távcső megnyomhatja a szemet, a könyv a fejünkre eshet vagy félrevezethet minket, amit látunk vagy

bejáró csomagkapcsolt adattovábbítási elvhez szükséges fejlesztések erőforrásainak előteremtését támogató egyik érvként használták a háborús helyzetben is épen maradó átviteli csatornákat.) Még a kilencvenes évek közepi, WWW-vel, grafikus böngészővel, e-maillal induló szakasz is akadémiai alapokra épült, a *gazdasági érdek és az üzleti innováció (profit-központú) teleológiája* csak az évtized végére, az úgynevezett dotcom-buborék és annak kipukkanása időszakára hódította el a technológiai oldal fejlesztését, amelyhez sokféle módon asszisztált a *stratégiai-politikai teleológia*, az ismert kulcsszavaival: demokratizálás, egyenlőség, hozzáférés, életminőség, versenyképesség, polgárbarát(abb)ság, olcsóbb működés. Azt, hogy az Internet olyanná lett, amilyen, az egymásra épülő teleológiai és technológiai korszakoknak köszönheti: mai tudásunk birtokában technológiailag is más módon fognánk a fejlesztésbe, szüntetnénk meg a szűk keresztmetszeteket, és az egyes teleológiák mentén is feltehetően önálló Internetek épülnének (külön akadémiai-tudományos, tartalomfogyasztási, üzleti és közösségi dimenzióban, ahol minden esetben másképp paramétereződnek a keresési, tárolási, kapcsolódási és fenntarthatósági szempontok).

⁶³ Latour „nem emberek” alatt – leegyszerűsítve – dolgok és szubjektumok hibridjeit érti, amelyekkel a cselekvések tere három eleművé válik.

⁶⁴ http://futureoflife.org/AI/open_letter

⁶⁵ http://futureoflife.org/static/data/documents/research_priorities.pdf

olvasunk, úgy a mesterséges intelligencia-megoldásaink sem kockázatmentesek. Egészen széles körben kelt izgalmat egy-egy új alkalmazás, és ettől szinte elválaszthatatlanul az automatizáció újabb hullámával elvesző munkahelyek kérdése. A környezetünkben egyre nagyobb számban jelenlévő és egyre több cselekvés révén életünk részévé váló intelligens robotokkal kapcsolatos kockázatkezelési, illetve erkölcsi és jogi dilemmák, s nem utolsósorban a megéltélenkülő párbeszéd a „kiborggá” válás veszélyeiről. Befejezésül röviden ezekről kell szót ejtenünk, annál is inkább, mert a filozófiai fogantatású technológiakritika mellett ezek a praktikus kérdések is rendre strukturálatlanul köszönnek vissza az alarmista megközelítésekben, ahogy arra a bevezető rész végén már nyomatékosan utaltunk.

AI és társadalom(tudomány): frontvonalak és viták

A nyugatószerkezet adminisztrációját már egy gép kezeli az egyik seattle-i kórházban. A Szilícium-völgy egyik szállodájában géplondiner viszi a törölközőt vagy az italt a vendégek szobáiba. A Los Angeles Times olyan cikket közölt egy földrengésről, amelyet egy szoftver írt. (Ezt az eredményeket szigorú szintaxissal közlő sport-sajtó már korábban megtette). A thai konyha ízeinek autenticitását, megfelelő fűszerezettségét robot méri 2014 óta, egy kormányprojekt eredményeképp. Watson, az IBM kvíznyertes bajnoka pedig már egészen másra is használja félelmetes szemantikus memóriáját: háborús veteránoknak ad tanácsot biztosítványválasztásban és életvezetési döntésekben, új receptek alkotásában segít séfeknek, vagy diagnózis felállításában orvosoknak.

A felsorolt friss példákat összegereblyező Miller (2014) a fentiekkel két állítást kíván illusztrálni:

- hogy a mesterséges intelligenciához köthető automatizáció korábban gyári és irodai munkahelyeket veszélyeztetett, napjainkban viszont már a tudásmunka és a szolgáltatások világába hatol be,
- s hogy többek között ez az oka annak, hogy míg korábban lehetett tudni, hogy a megszűnő állásokat a technológiai fejlődés újak teremtésével ellensúlyozza, addig ma már ez korántsem biztos.

Pontosabban kell azonban fogalmaznunk. Napjainkban még mindig párhuzamosan zajlik a fizikai⁶⁶ és a szellemi munka gépesítése, illetve automatizációja. Nem a gazdasági ágazat számít azonban, hanem a *végzett munka jellege*. Miller valamennyi példája a *repetitív agymunka* kiváltásáról szól. Ha nem a munkanélküliség lenne a kontextus, ennek az emancipatorikus és egalizáló jellegét hangsúlyoznánk inkább, hiszen ami „gépies”, az embertelen, a gépesítés így valójában humanizálás, akkor és amennyiben az életidő magasabb érték-hozzáadású, kreatív, nem automatizálható munkára szabadul fel. Az alacsony hozzá-

⁶⁶ Nagy robotizációs hullámra lehet számítani például azokban az országokban, ahol eddig a nagytömegű olcsó munkaerő volt a versenyképesség alapja (Kína-szindróma). Az automatizáció újabb hulláma elér eddig kevésbé érintett munkakör-kategóriákat (például takarítás, épületfenntartás, hulladék-kezelés). A legújabb technológiák némelyike pedig tudás-intenzívebbé tesz eddig kisiparinak és kézművesnek megmaradt területeket (például a 3D-nyomtatás az idomtermékeket előállító műhelyekben).

adást igénylő szakmák eltűnésével pedig közelebb kerülnek egymáshoz a munkaerőpiac különböző szegmensei. Brynjolfsson és McAfee (2014) azt gondolta végig, hogy mivel járhat ez azok oldaláról, akik kedvezményezettjei a változásoknak. Nekik személyes technológiák sora áll csatasorba, hogy képességeiket a leginkább megfelelő módon használhassák, és ezek mögött különösen fejlett infrastruktúra dübörög. Korlátlaná válik a hozzáférés a kultúrjavakhoz, és a nyersanyagok szerepét a szűk erőforrássá lett gondolatok és azok hordozói töltik be.

Azok számára azonban, akik rövid vagy középtávon kimaradnak mindebből, a jelenlegi irányítási rend és elosztási logika kétségkívül nem ígéri, hogy a profit-elvvel szemben az értéktermelésen alapuló munkahely-biztosítás kerülhetne az előtérbe. De ez nem technológiai, hanem kormányzási (governance) probléma. *A teljes gazdasági és társadalmi rend, az elosztás alapvető újrendezése* nélkül nem megoldhatóak az automatizálás-keltette munkaerőpiaci zavarok (Ford, 2015; Wadwha, 2015b). Ez pedig – hangsúlyozzák az elemzők – csakis az államok (újabbán: egyes városok (!)⁶⁷ újraértékelt szerepvállalásával, a köz-szféra kiterjesztésével működhet. Mindezt nagyban segítheti, ha az oktatás az új és nem a régi gazdaságra készít fel, a politika pedig nyitottá válik az új utak keresésére.

Nem árt tudatosítani, hogy nincs semmi szingularitás-specifikus a diskurzusban. Szuperintelligens gépek nélkül is érvényesül a mélységben erősen tagolt⁶⁸ automatizációs nyomás: az információtechnológiai forradalom „nem áll le”, a termelékenységet folyamatosan javítja (Byrne, 2013). Ha azonban a szuperintelligenciát a korábbiakban kifejtetteknek megfelelően „kivesszük a képletből”, akkor a munkahelyük megszűnésével riogató „csúcs-értelmiségiek” (orvosok, tanárok, jogászok, mesterszakácsok, sőt művészek)⁶⁹ számára sokkal inkább az új koegzisztenciák keresése és a munkájukat megkönnyítő, felgyorsító alkalmazások adaptálása lesz a kihívás, nem a székük elvesztése miatti aggodalom (Susskind és Susskind, 2016). Riogathat a hosszú távú gyötrelmet ígérő Sachs és Kotlikoff (2012), a munkanélküliek világát vizionáló Ford (2015), vagy a már egyenesen a munka világának végéről beszélő Smith (2013) – amennyiben a gazdasági alapok biztosíthatóak, *az oktatás, a tudomány, a kultúra és a humán szolgáltatások világa* korlátlanul felvevőképes lehet, s felszívhatja a (szükségszerűen egyre képzetesebb) munkaerőt.

S valóban: azokkal az előrejelzésekkel szemben, amelyek kizárólag a megszűnő munkahelyek számával, mennyiségbecslésével foglalkoznak, s amelynek alarmista részét találóan nevezi Marc Andreessen „luddita megtévesztésnek” (*Luddite Fallacy*), szaporodnak azok a tanulmányok, amelyek történeti kontextusban, a teljes munkaerőpiac méretének és szerkezetének számbavételével tesznek állításokat a technológiának a munkaerőpiacra gyakorolt hatásával kapcsolatban. Dajkó (2015) két friss szakanyagot ismertet, amelyek

⁶⁷ A feltétel nélküli alapjövdelemmel GMI (Guaranteed Minimum Income) való eredményes kísérletezés, egy fenntartható modell megtalálása például megfelelő válasz lenne, és a kérdéseket más tartományba tolná át. Újabbán nemcsak kormányok, hanem városi önkormányzatok is elkezdtek fontolgatni ezt a megoldást.

⁶⁸ Van, ahol kevesebben látják el a feladatokat, jobb információs infrastruktúrával. Van, ahol a termelés értékláncából esnek ki szereplők (és a munkahelyek a cégekkel együtt szűnnek meg vagy helyeződnek át máshová), és van, ahol egész iparágak tűnnek el.

⁶⁹ És evvel párhuzamosan a csodálatos verset író, gyönyörű képet festő, Mozartot lefőző robotok diskurzusát (legmerészebben: Levy, 2005) is ideje volna ad acta tenni.

végkövetkeztetése az, hogy „*a technológiai fejlesztések mindeddig több munkahelyet teremtettek, mint amennyit megszüntettek*”. A neves Forrester elemzője, J. P. Gownder a robotmunka döbbenetes előretörését prognosztizáló tanulmányok túlzásaira figyelmeztet, amelyek nem veszik figyelembe a nem pótolható emberi intelligenciát, majd (a düsseldorfi repülőtér példájával) érvel amellett, hogy a robotok és a mesterséges intelligencia alkalmazása a fejlettebb rendszerekhez igazodó összetettebb tudásokra támaszt folyamatos keresletet – tehát nem a kevesebb, hanem a képzetesebb munkavállaló a jövőképe.⁷⁰ A szintén mértékadónak számító Deloitte tanulmánya az utóbbi 150 évet Anglia és Wales példáján vizsgálva arra jutott, hogy „*bár a technológiai fejlesztések miatt a vizsgált időszakban nagy számban szűntek meg munkahelyek a mezőgazdaságban és gyártó iparágakban, mindezt bőségesen ellensúlyozta, hogy számtalan munkahely teremtődött az egészségügyben, a kreatív szektorokban, a technológiai szegmensben és az üzleti szolgáltatások területén*”. S itt annak ellenére nő szakadatlanul és magas százalékokkal kifejezhető módon a foglalkoztatottak száma, hogy az itt végzett munka egyre inkább technológia-intenzív és egyre hatékonyabb.

A fárasztó és monoton munkák eltűnése⁷¹ tehát együtt jár az oktatás különböző színterein alkalmazottak, a nővérek, gondozók, a közösségi szolgáltatásokban dolgozók (és tegyük hozzá: a tudományban, a médiában foglalkoztatottak vagy művészetből élő) számának gyarapodásával. S ha van alapja a Lanier-féle „új humanizmusnak”, azt sokkal inkább ebben a mozzanatban, és nem a mesterséges intelligencia elutasításában kellene keresni.

Nagy felbontásban azonban kiélesednek a valódi problémák is. Különböző munkakörökre és ágazatokra különböző mértékben lesznek igazak az átfogó trendek. Jól azonosítható, hogy a várakozásokkal ellentétben nem az alacsony, hanem a középszintű (és közepesen fizető) foglalkozások és állások tűnnek el (Autor és Dorn, 2013). Az elsődleges kihívás tehát a jövedelmi, vagyoni és társadalmi egyenlőtlenségek növekedése. Az átmenet

⁷⁰ A robotok alkalmazásának köszönhetően a német autóiparban 2010 és 2013 között az ágazat foglalkoztatottsága több mint 7 százalékkal nőtt, ekkor 10 000 alkalmazottra 1100 robot jutott. A Metra Martech által készített kutatás szerint a jelenleg működő egymilliónyi ipari robotnak csaknem hárommillió új munkahely köszönhető. Szakértőik megerősítik: a robotok nemcsak több, de jobb munkahelyet is teremtenek: „Az emberek szívesen dolgoznak robotokkal, mert megkönnyítik a munkájukat.” <http://nol.hu/tud-tech/yumi-a-robot-lebontotta-a-keritest-1574625> (Letöltve: 2015. december 20.)

⁷¹ Érdemes itt egy pillanatra elmélázni azon, mennyire biztosak az alarmisták abban, hogy a vezető nélküli járművek irtózatossá pusztítást végeznek majd a sofőrök sok tízmillió munkaadóján. Wadwha (2015a). S noha már van példa kötött pályás közlekedésben (metró, vonat) és a légi közlekedésben (robotpilóta, számítógéppel támogatott fel-és leszállás stb.) vezető nélküli megoldásokra, a mesterséges intelligencia-megoldások gyors bevezethetőségében hívők nem veszik figyelembe az elképesztő méretű cselekvőhálózatot, amelynek nemcsak az utak, az útjelzések, az eltérő közlekedési funkciók, a különböző úti célok és minden egyes különböző közlekedő ember a része, s ami a változást nem a technológiai előrehaladás, hanem sokkal összetettebb paraméterek függvényévé teszi. De eközben azt is látni kell, hogy a sofőrmunka önmagában monoton, alacsony értékhozzáadású munka, amelynek lecserélése indokolt – de nem biztos, hogy pusztán a vezető nélküliség az egyetlen célfüggvény. Egy új funkcionális térben a szükséges közlekedési teljesítmény válhat csökkenhetővé, vagy a közlekedés menedzsmentje tudás-intenzív (ahogyan például speciális munkagépek – daruk, kombájnok stb. – esetében látjuk: ezekhez ma jellemzően diplomás, nagy tudású kezelők-vezetők kellenek).

kétségkívül lehet fájdalmas, egyenesen brutális, ahogy Kaplan (2015) fogalmaz, ha tőke és munka csatájában a tőke felé billen a mérleg. Ha a bérek növekedése tartósan alulmarad a termelékenység növekedése mögött. Ha a munkaerőpiac nem rugalmasabbá, hanem merevebbé válik. Ezek a kérdések azonban már mesterséges intelligencia és automatizáció-kontextus nélkül is érvényes és gazdagon tárgyalt makroökonómiai alapdilemmák. Bernstein (2015) szerint a sarkalatos kérdés az, hogy növekszik-e a sebessége az emberi munka technológiai kiváltásának – és válasza az, hogy *nem* (amelyre számtalan bizonyítékot sorol fel).

És vajon az emberi testben mit válthatnak ki/javíthatnak meg a mesterséges intelligencia jelenlegi rendszerei, milyen élethelyzetben és milyen célcsoportoknál? A *biológiai test korlátainak leküzdésében* szerepet játszó megoldások valódi és érvényes tartalommal töltik fel az 1960-ban született cyborg-narratívát.⁷² Humanistának látjuk azokat az erőfeszítéseket, amelyekkel fogyatékkal élő és beteg embertársaink jutnak lehetőséghez egy méltóságtelesebb, autonómabb életre (látás-hallás- és mozgásprotézisekkel, de akár gondolat-vezérléses környezet-manipulációval⁷³). Az exoskeleton, a külső váz használata kíméli a nagy súlyt emelgető kórházi ápolókat és munkásokat. Azok az augmentált valóság-rendszerek (ahol a műveletvégzést mesterséges intelligencia-megoldások támogatják) és azok a környezetek, ahol testünk gépi kiterjesztését sajátunkként érzékeljük (s amit Boyd (2015) szuperpropriocepciónak nevez), a lehetőségek egy *speciális* világát jelentik. A szervezet egészségi állapotának megőrzése, monitoringja és gyógyítása, vagy a minőségi öregedés elősegítése érdekében latba vetett *általános, mindenkit érintő* technológiák pedig együttesen azt ígérik, hogy az evolúciós „előtörténetre” a gépi intelligencia felhasználásával épül új „réteg”, amely nemcsak az egyes embereknek kínál jobb-létet, hanem a fajnak is nagyobb esélyt a túlélésre (Naam, 2005).

A test „megerősítésének” (enhancement) tudományos programjával kapcsolatos lényegi viták az ezredforduló után már lezajlottak,⁷⁴ napjainkban „középutas” álláspontok jelennek meg: Buchanan (2011) megközelítése „tervezési hibaként” szeretné láttatni a feljavítandó sajátosságokat, Agar (2013) pedig az ember „enyhe felerősítésének” (*moderate human enhancement*) programja mellett tör lándzsát, mert szerinte a túlzásba vitt változat az egész emberi identitást ásná alá. Abban azonban szinte mindenki egyetért, hogy megfelelő keretek között tartva a szuperintelligencia helyett mindez egy „új szuperhumán korszak” (new superhuman age) eljövételét gyorsíthatja fel (Boyd, 2015).

⁷² Noha a diskurzus-indító szerzők, Clynes és Kline (1960) specifikus úrkutatási kontextusban fogalmazták meg az emberi testnek a Földön kívüli körülmények elviselését lehetővé tevő mérnöki „továbbfejlesztését”, a diskurzus mára sokkal összetettebb és általánosabb jellegű lett.

⁷³ Ennek a jövő interfészei szempontjából különösen fontos és izgalmas iránynak számos sikeres előzmény után 2014 őszén sikerült egymástól 8000 kilométerre lévő emberi agyak között sebészi beavatkozás nélküli kapcsolatot létesíteni. Az agy-számítógép kapcsolattal Franciaország és India között sikerült gondolatokat (például a kéz és a láb megmozdítására vonatkozó instrukciót) közvetíteni. http://index.hu/tudomany/2014/09/05/mar_mukodik_az_internetes_gondolatvitel/

⁷⁴ A pozitív ígéretek foglalataként megszületett a transzhumanizmus fogalma, amelyre válaszul megjelentek a bio-ludditák is, akik korlátozni szeretnének minden ezirányú fejlesztést (Young, 2005). Garreau (2006) számos foratókönyvet gondolt végig, Hughes (2004) a „demokratikus transzhumanizmus” jegyében az egész folyamat közösségi kontrolljának szükségességére figyelmeztetett. Az elmúlt időszak termésére lásd Savulescu és Bostrom (2011) olvasókönyvét és Anissimov (2015) mérész vízióját, amelyben a szuperintelligenciát és a nanotechnológiát kapcsolja össze (az „ember digitális feljavításaként”) felfogott transzhumanizmussal.

A civilizációs előrelépésnek ez az ígérete, ha megfelelően artikulált, ellensúlyozhatja azokat a valódi és kétségtelen veszélyeket, amelyek már a gépi intelligencia-megoldások jelenlegi világában is jól detektálhatóak. Régóta tudjuk, hogy az algoritmikus alapokon működő kereskedő-ágensek a tőzsdéken okozhatnak anomáliákat. Az automatizált és összekapcsolt nagy hálózati rendszerekben, a kritikus infrastruktúrákban a rossz cselekvésválasztás vagy az ellenséges támadás miatt alakulhatnak ki működési zavarok (például áramszünet, közlekedési káosz, tranzakciós rendszerek leállása stb.). A viselhető és perszonalizált okos eszközök és a felhő-architektúra megnöveli az intelligens hálózati bűnözésnek való „cyber-kitettséget”.⁷⁵ S mi történik, ha a bármely orvos-virtuóznál biztosabb kezű sebészrobot hibát követ el? Hogyan birkózhat meg egy rendkívüli helyzettel egy vezető nélküli jármű? Milyen szintű döntést hozhat gépi intelligencia, amelynek végén embereket érintő tranzakció áll?⁷⁶

Az ilyen és ehhez hasonló kérdések jogi és erkölcsi dilemmák sorát vetik fel, emiatt nemcsak a fejlesztőknek kell az efféle aggodalmakat kiemelten kezelniük. Avval, ahogyan a baleset vagy probléma forrását jelentő géptől egyre hosszabb láncokkal jutunk el a felelősséget vállalni képes emberekhez vagy intézményekhez, kezdeni kell valamit. A tervezési, a kivitelezési, a működtetési, a karbantartási, a tesztelési és a felhasználói érintettségek világát szituáció-érzékenyen kell tudni „telepíteni” és felosztani. *Teljesen téves irány azonban azt fontolgatni, miképpen tehetők jogi személylé, jogok birtokosává (és eképpen felelőssé) intelligens robotok* (Lin et al., 2014). Hasonlóképpen zsákutca az etikus robot, a morális döntéshozatalra alkalmas gép megteremtésének programja (Allen – Wallach, 2010). *A hibrid rendszerben az erkölcsi felelősség minden esetben az emberi komponensé. A gépek programjuknak megfelelően viselkednek. Nem állíthatóak erkölcsi kihívások elé, mert – mint korábban láttuk – már a jelentések országába sem bocsáttattak be, így aztán a különlegesen bonyolult, sokféleképpen referenciális metajelentések világa még inkább távol van tőlük.*⁷⁷ A mesterséges intelligencia ágenseinek nincs szuverén döntése, kizárólag

⁷⁵ A cyberbűnözéssel foglalkozó európai szakintézmény (European Cybercrime Center) 2014-es szakanyaga – *The Internet Organised Crime Threat Assessment* (iOCTA) – valamennyi területet áttekinti, és új, a fejlettebb technológiákhoz kötődő elkövetési módok megjelenését jósolja (<https://www.europol.europa.eu/content/internet-organised-crime-threat-assessment-iocta>). További példákra lásd: <http://www.origo.hu/techbazis/20141021-okoseszkoz-hacker-gyilkosság-implantatum-gyogyaszat.html>

⁷⁶ Vegyük észre: amikor sorszámot húzunk, ügýtípusok közül választva, egy primitív mesterséges intelligencia-megoldás „dönti el”, hogy mikor és melyik ügyintézőnél kerülünk sorra. Erre gyakorlatilag érzéketlenek vagyunk, mert az egyszerű besorolási algoritmus nem sért érzékenységet, és felül is írható (ha sietünk, az előbbre soroltaktól és az ügyintézőktől is kérhetjük a méltányosságot). Ott azonban, ahol – akár hasonlóan banális esetben – már nincs mód a beavatkozásra, már komoly aggályok vetődnek fel. Nem beszélve arról, amikor „buta” algoritmusok minősítenek egy életbiztosítást kockázatosabbnak pusztán a testsúly növekedése miatt, nem törődve avval, hogy az háj vagy izom-e (ami inkább csökkentené a kockázatot). Emiatt alakult ki sok alrendszerben az a gyakorlat, hogy az embereket érintő gépi operáció végpontján afféle „humán kontrollként” mindig a gépnél komplexebb körülmény-együttest kezelni képes emberek állnak, akik felülbírálnak/módosíthatják az automatizált folyamat végeredményét.

⁷⁷ Azok a próbálkozások, amelyek erkölcsi dilemmaként próbálják beállítani a döntéshelyzetbe kerülő gépek esetét (mint például a ’mit kezd a robot akkor, ha nem egy, hanem két emberpótlékot kellene megmentenie a gödörbe zuhanástól’, vagy ’mit tesz az automata sofőr, ha választani kellene, ki kerüljön el egy baleset: saját, kevés utassal közlekedő járműve vagy egy sok gyereket szállító busz’), megtevesztő paradiskurzusokhoz vezetnek (<http://sg.hu/cikkek/107784/robotok-az-etika-csapdajaban>). Ezek ugyanis *nem erkölcsi, hanem programozási problémák*, amelyek mögött nem új típusú, hanem teljesen *hagyományos etikai kérdőjelek és megfontolások* állnak.

előfeltételeknek és szabályoknak olyan együtteseivel bírnak, amelyet a rendszereket megalkotó emberek plántáltak beléjük. Emiatt nagyon fontos kérdés a mesterséges intelligencia pragmatikus etikai és átfogó morálfilozófiai kérdéseivel foglalkozni – csak nem a gépek oldaláról, hanem az emberekről.

Ahhoz például, hogy a jogrendszer kezelni legyen képes a praktikus problémákat, először is le kellene tudni fordítani a jog nyelvére a mesterséges intelligenciát. De hogyan definiálható kiindulásként maga az intelligencia fogalma, ráadásul az embertől függetlenül – ha annak négy kulcs-komponense, a tudatosság, a gondolat, a szabad akarat és az elme mibenléte Arisztotelész óta állandóan diszkutált tárgya a tudománynak?⁷⁸ Markus Hutter és Shane Legg használhatónak tűnő, rövid és szellemes definíciót alkotott. Náluk „*az intelligencia egy ágens képességét méri, amellyel céljait különböző környezetekben eléri*” (Lea, 2015), ahol az ágens egyaránt lehet ember és gép. Vegyük észre, hogy „a cél elérése” mozzanata szellemesen magába foglalja a tervezés, a tanulás és a problémamegoldás másutt önállóan tárgyalt attribútumait. De hol azonosítjuk az ágens a hibrid rendszerben? Továbbá: minden mérnöki alkotásban benne rejlik a lehetőség, hogy méretezése, hibátűrése olyan valóságos kihívással találkozik, amelyre a tervező nem lehetett felkészülve. Elfogadjuk ezt egy „okos gép” esetében is? Mi, emberek, gyakorta vállalunk kényszerűen kockázatot, mert döntés- és cselekvéskényszerben vagyunk információhiányos helyzetben is – de ebből mennyit tudunk ráterhelni a gépi oldalra? Amikor feladatokat delegálunk, a felelősség meddig marad a miénk, különös tekintettel a cselekvő-hálózati beágyazásra? És mi az, amit nem delegálunk? Lehet mindezt listázni egy alapidokumentumban (Havens, 2015)?

S még ha sikerülne is megnyugtató intelligencia-formuláig vagy AI etikai kódexig jutni, hogyan határozhatjuk majd meg például a „jó”-t (Armstrong, 2014), minden kontextusok egyik legfontosabbikát, amikor a hibrid rendszer céljait definiáljuk, és használatának hatáskövetkezményeire készülünk? S mindezt hogyan tudjuk lefordítani gépi kódra?

A kérdések még folytathatóak lennének – csak jó volna túljutni az alarmista útkadályokon, hogy valóban a lényeges területekre koncentrálhassunk. Bemelegítésnek megteszi az a 2015 végén megjelent válogatás, amely 175 kortárs tudós, filozófus és művész tollából tartalmaz olykor aforisztikus, olykor tanulmány-értékű válaszokat a fenti és sok más, a gondolkodó gépekkel kapcsolatos kérdésre (Brockman, 2015). Ebben egyszerre vannak jelen a keresett és sürgetett új szemlélet irányába mutató megközelítések és az alarmista manifesztumok – ám kézzelfogható közelségbe kerül a frissen megfogalmazott AI-ígéretekből való kezdődő kiábrándulás, az eddigi kilencet követő, közelítő tizedik AI-tél (AI-winter) is.⁷⁹

Pedig a mesterséges intelligencia valódi helye az „örök tavasz”. Technológia-alapúvá lett kultúránk egyik legfontosabb, legperspektivikusabb tudományos iránya, kutatás-fejlesztési és alkalmazási területe, amelyben minden kis előrelépés avval kecsegtet, hogy az ember-gép hibrid rendszerek hatékonyabban tudnak teljesíteni az élet minőségének javításában és a civilizációs kihívások kezelésében. Az alarmizmus legnagyobb bűne, hogy talmi szenzációkeltésével figyelmet és erőforrásokat von el a valódi diskurzusoktól.

⁷⁸ 2014 nyarán egy orosz milliárdos, Dmitrij Volkov kísérletet tett arra, hogy a jachtjára meghívott neves filozófusokkal és társadalomkutatókkal konszenzusig sikerüljön jutni a kérdésben. David Chalmers, az egyik résztvevő egyenesen azzal búcsúzott az eredménytelen „csúcstalálkozótól”, hogy a megoldásra a következő évszázadig kell majd várni.

⁷⁹ S mindeközben talán érdemes újra kézbe venni Darab Tamás könyvecskéjét (Darab, 1991), amely a mesterséges intelligencia körüli sarkalatos ismeretelméleti viták élvezetes és precíz összefoglalása, s amelynek köpönyegéből a kérdéskör hazai irodalma kinőtt.

Irodalom

- Agar, Nicholas 2013: *Truly Human Enhancement: A Philosophical Defense of Limits* The MIT Press
- Allen, Colin – Wallach, Wendell 2010: *Moral Machines: Teaching Robots Right from Wrong* Oxford University Press
- Anissimov, Michael 2015: *Our Accelerating Future: How Superintelligence, Nanotechnology, and Transhumanism Will Transform the Planet* Zenit Books,
- Armstrong, Stuart 2014: *Smarter Than Us: The Rise of Machine Intelligence*, Machine Intelligence Research Institute
- Autor, David H. – Dorn, David 2013: How Technology Wrecks the Middle Class *New York Times*, August 24. 24–27. http://www.collier.sts.vt.edu/eng14874/pdfs/autor_nyt_9_24_13.pdf
- Baranyi, Péter 2012: Kognitív infokommunikáció: egy ébredő interdiszciplína http://otodikal.program.huminf.u-szeged.hu/sites/default/files/BaranyiPeter_Coginfocom_szeged'12.pdf
- Barrat, James 2014: *Our Final Invention: Artificial Intelligence and the End of the Human Era* Thomas Dunne Books
- Basulto, Dominik 2015: The strange link between global climate change and the rise of the robots *The Washington Post*, September 8. <http://www.washingtonpost.com/news/innovations/wp/2015/09/08/the-strange-link-between-global-climate-change-and-the-rise-of-the-robots/>
- Berger, Viktor 2008: Bruno Latour tudományképe és antropológiai megközelítésmódja *Szociológiai Szemle* 4. 72-92.
- Bernstein, Jared 2015: *The Reconnection Agenda: Reuniting Growth and Prosperity* CreateSpace Independent Publishing Platform
- Blackford, Russell – Broderick, Damien 2014: *Intelligence Unbound: The Future of Uploaded and Machine Minds* Wiley-Blackwell
- Bodnar, Ken 2015a: Dawkins, Wasps, Artificial Intelligence, Evolution, Memorability and Artificial Consciousness *Future Imperfect & Software Stream of Consciousness* December <http://coderzen.blogspot.hu/2015/12/dawkins-wasps-artificial-intelligence.html>
- Bodnar, Ken 2015b: Dimension & Event Sorters & Classifiers - The Genesis of Artificial Consciousness & Abstract Machine Reasoning *Future Imperfect & Software Stream of Consciousness* December <http://coderzen.blogspot.hu/2015/12/dimension-event-sorters-classifiers.html>
- Bölcső, Dániel 2015: Robot ölt embert, ez már az apokalipszis? *Index*, Július 7. http://index.hu/tech/2015/07/08/robot_gyilkosság_buntetojog_gepek_vilagvege/
- Bostrom, Nick 2014: *Superintelligence: Paths, Dangers, Strategies* Oxford University Press
- Boyd, Richard 2015: Man Vs. Machine: How Humans Are Driving The Next Age Of Machine Learning *Techcrunch*, June 15. aug. 18. <http://techcrunch.com/2015/06/11/man-vs-machine-how-humans-are-driving-the-next-age-of-machine-learning/#.pzubov:ck1O>
- Brain, Marshall 2015: *The Second Intelligent Species: How Humans Will Become as Irrelevant as Cockroaches* BYG Publishing
- Brockman, John 2015: *What to Think About Machines That Think: Today's Leading Thinkers on the Age of Machine Intelligence* Harper Perennial
- Brooks, Rodney, 2003: *Flesh and Machines: How Robots Will Change Us* Vintage
- Brynjolfsson, Erik – McAfee, Andrew 2014: *The Second Machine Age. Work, Progress and Prosperity in the Second Machine Age* W. W. Norton & Company;
- Buchanan, Allen 2011: *Better than Human: The Promise and Perils of Enhancing Ourselves* Oxford University Press
- Burgess, Mark 2015: *In Search of Certainty. The Science of Our Information Infrastructure* O'Reilly Media
- Byrne, David M. – Oliner, Stephen D. – Sichel, Daniel E. (2013) Is the Information Technology Revolution Over? *Staff Working Paper*, FEDS March (No.3). <http://www.federalreserve.gov/pubs/feds/2013/201336/201336pap.pdf>

- Callon, Michel 2005: Alakuló társadalom. A technika, mint a szociológiai elemzés eszköze *Replika* 51-51. (november) 89-105.o.
- Card, Stuart K. – Moran Thomas P. – Newell, Allen 1983: *The Psychology of Human-Computer Interaction* Erlbaum, Hillsdale
- Clynes, Manfred E. – Kline, Nathan S. 1960: Cyborgs and Space *Astronautics* September 24-26, 74-76. <http://cyberneticzoo.com/wp-content/uploads/2012/01/cyborgs-Astronautics-sep1960.pdf>
- Colvin, Geoff 2015: *Humans Are Underrated: What High Achievers Know That Brilliant Machines Never Will* Portfolio
- Dajkó Pál 2015: A robotok több munkahelyet teremtenek, mint amennyit elvesznek *IT café* augusztus 28. http://itcafe.hu/hir/robot_tecnologia_munkahely.html
- Darab Tamás 1991: *A gépesített értelem* Áron Kiadó
- Del Monte, Louis A. 2014: *The Artificial Intelligence Revolution: Will Artificial Intelligence Serve Us Or Replace Us?* Amazon Kindle Edition
- Diebold, John 1969: *Man and the Computer: Technology as An Agent of Social Change* Praeger
- Domingos, Pedro 2015: *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* Basic Books, 2015
- Dreyfus, Hubert 1972: *What computers can't do: A Critique of Artificial Reason* New York: Harper & Row
- Dreyfus, Hubert 1992: *What computers still can't do": A Critique of Artificial Reason.* Cambridge, MA: MIT Press
- DuBravac, Shawn 2015: *Digital Destiny: How the New Age of Data Will Change the Way We Live, Work, and Communicate* Regnery Publishing
- Engelbart, Douglas 1962: *Augmenting Human Intellect* paper, October <http://www.dougenelbart.org/pubs/augment-3906.html>
- Floridi, Luciano: *The Fourth Revolution: How the infosphere is reshaping human reality* Oxford University Press, 2014
- Ford, Martin 2015: *Rise of the Robots: Technology and the Threat of a Jobless Future.* Basic Books
- Garreau, Joel 2006: *Radical Evolution: The Promise and Peril of Enhancing Our Minds, Our Bodies—And What It Means to Be Human* Broadway Books
- Gitt, Werner 2004: *Kezdetben volt az információ* Evangéliumi Kiadó (2., bővített és javított kiadás)
- Gleiser, Marcelo: Sinister dreams of transhuman machines: or, the world as information In: Uő: *The Island of Knowledge. The Limits of Science and the Search for Meaning* (31. fejezet) Basic Books, 2014
- Goertzel, Ben 2014: *Ten Years To the Singularity If We Really Really Try: ... and other Essays on AGI and its Implications* CreateSpace Independent Publishing Platform
- Good, Irving John 1965: Speculations Concerning the First Ultraintelligent Machine *Advances in Computers*, Vol. 6.
- Grier, David Alan 2005: *When Computers Were Human* Princeton University Press
- Grove Andrew S. 1998: Csak a paranoidok maradnak fenn Bagolyvár Kiadó
- Grudin, Jonathan 2007: A moving target: The evolution of human-computer interaction. In Sears Andrew – Jacko Julie A. (Eds.): *Human-Computer Interaction Handbook* (2nd Edition). CRC Press
- Havasi, Catherine 2014: Who's Doing Common-Sense Reasoning And Why It Matters *TechCrunch*, Augusztus 9. <http://techcrunch.com/2014/08/09/guide-to-common-sense-reasoning-whos-doing-it-and-why-it-matters/>
- Havens, John C. 2015: The importance of human innovation in A.I. ethics *Mashable*, October 3. <http://mashable.com/2015/10/03/ethics-artificial-intelligence/?curator=MediaREDEF#9W0.aNQo7gqc>
- Hernandez, Daniela 2016: It's hard work being funny – especially for robots *Fusion*, January 8. <http://fusion.net/story/251798/funny-robots/>
- Horváth Bence 2015: Aggódní azon, hogy a robotok ellenünk fordulnak, olyan mintha a Mars túlnépesedésén aggódnánk *444.hu* május 25., hétfő <http://444.hu/2015/05/25/aggodni-azon-hogy-a-robotok-ellenunk-fordulnak-olyan-mintha-a-mars-tulnepesedesen-aggodnank/>

- Hsu, Stephen 2015: Don't Worry, Smart Machines Will Take Us With Them. Why human intelligence and AI will co-evolve *Nautilus*, September 3. <http://nautil.us/issue/28/2050/dont-worry-smart-machines-will-take-us-with-them>
- Hughes James 2004: *Citizen Cyborg: Why Democratic Societies Must Respond To The Redesigned Human Of The Future* Basic Books
- Kaplan, Jerry 2015: *Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence* Yale University Press
- Király Gábor 2005: Hovatovább STS? Fejtegetések az értelmezési flexibilitás, a hiányzó tömeg, a kiborg és a demokrácia kapcsán *Replika* 51-52. November 25-56.o.
- Knight, Will 2015: Can This Man Make AI More Human? MIT Technology Review December 17. http://www.technologyreview.com/featuredstory/544606/can-this-man-make-ai-more-human/?utm_campaign=newsletters&utm_source=newsletter-weekly-computing&utm_medium=email&utm_content=20151217
- Kreft, Hans-Dietrich, 2003: *Humatics – Theorie Der Operablen Wissenseigenschaften. Geld und Wissen* Weissensee-Verlag, Berlin
- Kurzweil, Ray 2013: *A szingularitás küszöbén* Ad Astra (eredetije: *The Singularity is Near*. New York: Viking Books, 2005)
- Kurzweil, Ray 2013: *How to Create a Mind* Penguin Books
- Lea, Gary 2015: Why we need a legal definition of artificial intelligence *World Economic Forum Agenda*, September 7. https://agenda.weforum.org/2015/09/why-we-need-a-legal-definition-of-artificial-intelligence/?utm_content=buffer4f517&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer
- Levy, David 2005 *Robots Unlimited: Life in a Virtual Age* A K Peters/CRC Press
- Levy, David 2008: *Love and Sex with Robots: The Evolution of Human-Robot Relationships*
- Licklider, Joseph Carl Robnett (1965): *Libraries of the Future* Cambridge, MA
- Licklider, Joseph Carl Robnett (1960): "Man-Computer Symbiosis" *IRE Transactions on Human Factors in Electronics*, vol. HFE-1, 4-11, March 1960.
- Lin Patrick – Abney Keith – Bekey George A.: 2014: *Robot Ethics: The Ethical and Social Implications of Robotics* The MIT Press
- Littman, Michael 2015: Rise of the Machines' is Not a Likely Future *Live Science*, January 28. <http://www.livescience.com/49625-robots-will-not-conquer-humanity.html>
- Harper Perennial
- Love, Dylan 2014: By 2045 'The Top Species Will No Longer Be Humans,' And That Could Be A Problem *Business Insider* July 6. <http://www.businessinsider.com/louis-del-monte-interview-on-the-singularity-2014-7#ixzz3izPDTrft>
- Markoff, John 2015: *Machines of Loving Grace. The Quest for Common Ground Between Humans and Robots* Ecco
- Meier, Patrick 2015: *Digital Humanitarians How Big Data is Changing the Face of Humanitarian Response* CRC Press
- Michelucci, Pietro - Dickinson, Janis L. 2016: The power of crowds *Science*, 2016 January 32-33.o.
- Miller, Claire Cain 2014: As Robots Grow Smarter, American Workers Struggle to Keep Up *The New York Times* The Upshot Dec.15. http://www.nytimes.com/2014/12/16/upshot/as-robots-grow-smarter-american-workers-struggle-to-keep-up.html?utm_source=nextdraft&utm_medium=email&abt=0002&abg=1&r=2
- Miller, James D. 2012: *Singularity Rising: Surviving and Thriving in a Smarter, Richer, and More Dangerous World* BenBella Books,
- Mindell, David A. 2015: *Our Robots, Ourselves: Robotics and the Myths of Autonomy* Viking
- Moravec, Hans 1988: *Mind Children* Harvard University Press
- Moravec, Hans 1998: *Robot: Mere Machine to Transcendent Mind* Oxford University Press

- Muehlhauser, Luke 2013: *Facing the Intelligence Explosion* Machine Intelligence Research Institute
- Naam, Ramez 2005: *More Than Human: Embracing the Promise of Biological Enhancement* Broadway Books
- O'Leary, Katie – Wobbrock, Jacob O. – Riskin, Eva A. (2013): Q-Methodology as a Research and Design Tool for HCI CHI 2013 April 27-May 2, Paris, France. <http://students.washington.edu/kathlo/HCI-Q-CAMERARReady.pdf>
- Parkin, Simon 2015: The Brain in the Machine *How We Get to Next*, October 21. <https://howwegettonext.com/the-brain-in-the-machine-f61559a38887#.gji39ffcw>
- Pesthy, Gábor 2015: AI: a mesterséges intelligencia barát vagy ellenség *Origo*, 2014. Jan. 4. <http://www.origo.hu/tudomany/20141222-ai-a-mesterseges-intelligencia-barat-vagy-ellenseg.html>
- Preece, Jenny – Sharp, Helen, Rogers, Yvonne 2015: *Interaction Design: Beyond Human-Computer Interaction* 4th ed. John Wiley & Sons
- Rees, Martin 2004: *Our Final Century: The 50/50 Threat to Humanity's Survival* Arrow Books
- Rothblatt, Martine 2014: *Virtually Human: The Promise - and the Peril - of Digital Immortality* St. Martin's Press
- Rovenszkij – Ujemov – Ujemova 1964: A gép és a gondolkodás Kossuth Könyvkiadó, Budapest
- Savulescu, Julian – Bostrom, Nick 2011: *Human Enhancement* Oxford University Press
- Sachs, Jeffrey – Kotlikoff, Laurence J. 2012: Smart Machines and Long-Term Misery *NBER Working Paper* No. 18629, December <http://www.nber.org/papers/w18629>
- Smith, Noah 2013: The End of Labor: How to Protect Workers From the Rise of Robots *The Atlantic*, January 14. <http://www.theatlantic.com/business/archive/2013/01/the-end-of-labor-how-to-protect-workers-from-the-rise-of-robots/267135/>
- Stone, Maddie 2014: The Dominant Life Form in the Cosmos Is Probably Superintelligent Robots *Motherboard*, dec.19. <http://motherboard.vice.com/read/the-dominant-life-form-in-the-cosmos-is-probably-superintelligent-robots>
- Storm, Benjamin C. – Stone, Sean M. 2015: Saving-Enhanced Memory. The Benefits of Saving on the Learning and Remembering of New Information *Psychological Science* February Vol. 26 No. 2. 182-188
- Susskind Richard – Susskind, Daniel 2016: *The Future of the Professions: How Technology Will Transform the Work of Human Experts* Oxford University Press
- Thackara, John (2005): *In the Bubble: Designing in a Complex World* Cambridge, MA, and London: MIT Press
- Toyama, Kentaro 2015: *Geek Heresy. Rescuing Social Change from the Cult of Technology* Public Affairs
- O'Callaghan, Jonathan 2015: This Robotic Assassin Will Hunt And Kill Starfish That Are Destroying The Great Barrier Reef *IFL Science*, September 3. <http://www.iflscience.com/technology/robotic-assassinator-will-hunt-and-kill-coral-reef-destroying-starfish>
- Urban, Tim 2015: The AI Revolution: The Road to Superintelligence *Wait But Why* 1. <http://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html>
- Wadwha, Wiwek 2015a: It's No Myth: Robots and Artificial Intelligence Will Erase Jobs in Nearly Every Industry *SingularityHub* 2015. július 7. <http://singularityhub.com/2015/07/07/its-no-myth-robots-and-artificial-intelligence-will-erase-jobs-in-nearly-every-industry/>
- Wadwha, Wiwek 2015b: Should We Redesign Capitalism to Address Our Jobless Future? *SingularityHub*, 2015. július 20. <http://singularityhub.com/2015/07/20/we-need-a-new-version-of-capitalism-for-the-jobless-future/>

Young, Simon 2005: *Designer Evolution: A Transhumanist Manifesto* Prometheus Books

Z. Karvalics, László – Juhász, Lilla 2008: *Társadalmi informatika I.* Információs Társadalomért Alapítvány, Budapest

Z. Karvalics, László 2015: Minden dolgok Internetje (Internet of Everything) In: Z. Karvalics László (szerk.): *Metszéspontok. Társadalomtudomány és infokommunikáció az ezredforduló után* Gondolat/Infonia, Budapest, 216-246.o.

