

ÖSSZEGZÉS
(KISS KÁROLY)

Individualizáció és személyiség – történelmi áttekintés

A nyugati civilizáció kiemelkedésében, csodálatos felívelésében az *individualizáció* kulcsfontosságú szerepet játszott. Az egyén, a személyiség a világ semelyik más részén nem fejlődött ki ilyen mértékben, nem jutott ilyen fontos szerephez. Nem véletlen, hogy a keleti társadalmak máig nem tudtak felzárkózni a Nyugathoz, kivéve azokat, ahol az individualizmusra épülő kapitalizmus vált uralkodóvá a korábbi ideológiákkal és vallásokkal szemben.¹

A két és fél ezer éves nyugati civilizáció bölcséleti alapja az arisztotelészi dualizmus, az objektum és a szubjektum szétválasztása. Ebből következik a nyugati kétértékű logika és ontológia is. Ezen az úton haladva jutunk el az egyén, a személyiség mai felfogásához. Gazdasági oldalról a magántulajdon, szellemileg pedig a kereszténység (Isten megtestesülése Jézusban) alapozta meg a nyugati személyiségfelfogást.

Az összefoglaló táblázatban a különféle vallásokat, filozófiákat és ideológiákat aszerint csoportosítjuk, hogy meddig jutottak el a személyiség fejlődésének útján. Hat lépcsőfokot különböztetünk meg:

- nincs önálló *én*, a személyiség a világszellem része;
- a világ szétvált objektumra és szubjektumra, a lélek önálló entitás;
- az ember áll a világmindenség középpontjában;
- csak az *én* létezik, a világ általa teremtett illúzió;
- a tudat az agy terméke;
- az *én* nem egységes, komponensekből áll.

¹ A „fejlődés” fogalmát a szokásos értelemben használom. Ugyanakkor látnunk kell, hogy a nyugati „fejlődés” okozta a környezeti válságot, bolygónk élővilágának – és talán megszokott klímájának is – a tönkretételét. A keleti vallások és filozófiák talaján ez nem következett volna be.

A személyiség a különböző vallásokban, filozófiákban és ideológiákban

	jellemzők					
vallások, filozófiák, ideológiák (időrendben)	nincs önálló <i>én</i> , a személyiség a világgal szemben	a világ szétvárt objektumra és szubjektumra, a lélek önálló entitás	az ember áll a világmindenség középpontjában	csak az <i>én</i> létezik, a világ általa teremtetett illúzió	a tudat az egy terméke	az <i>én</i> nem egységes, komponensekből áll
zsidó vallás		elismeri a lélek létezését, a szadduceusokat kivéve				
taoizmus (a <i>Wu</i> tudás)	az <i>én</i> a természet közös felsőbb ÉN-jének átélése az ember az univerzum része					
konfucianus filozófia	az <i>átman</i> nak, az egyéni léleknek van kozmikus szintű, univerzális megjelenése is, amely lényegét tekintve azonos a <i>brahman</i> nal (a világszellemmel); azonos vagy a <i>brahman</i> nal, és ezáltal mindennel és mindenkivel	az <i>átman</i> az egyéni lélek				
<i>védánta</i> bölcselet (hindu vallás)		létezik egy magasabb rendű, anyagtalan, megiszult létforma, magasabb rendű <i>én</i> (<i>átman</i> , <i>purusa</i>), aki újra-születések sorozatán keresztül ismételtén visszakerül az anyagi világba				
<i>upanisadok</i> (a hindu vallás tanításai)		minden élőlényhez tartozik egy-egy <i>purusa</i> ; ez az ő valódi önzonosság, felsőbb ÉN-je				
<i>számhija</i> bölcselet (hindu vallás)						
buddhizmus (1)						

		jellemzők			
		a világ szétvált objektumra és szubjektumra, a lélek önálló entitás	csak az <i>én</i> létezik, a világ általa teremtet illúzió	a tudat az egy terméke	az <i>én</i> nem egységes, komponensekből áll
vallások, filozófiák, ideológiák (időrendben)					
Arisztotelész előtti ókori görög filozófiák	a lélek az újraszületések után visszatér az istenek honába (rokonság a távolkeleti filozófiákkal)	a lélek halhatatlan és újra-születések sorozatán megy keresztül			
Arisztotelész		dualista filozófia			
kereszténység		Isten megtestesült Jézusban, személyiséggé vált			
iszlám		a jó lelkek a Paradicsomba jutnak			
reformáció, protestantizmus		üdvözülés a munkán keresztül, kiválasztottság, hivatástudat			
klasszikus német idealizmus	van egy <i>transzcendentális szubjektum</i> (a fichtei Abszolút Én, a schellingi Abszolútum, a hegeli Szellem), amelyben az objektum és a szubjektum egybeesik, amely szétoszja önmagát szubjektumra és objektumra és amely önmagán keresztül közvetít e két pólus között	a <i>transzcendentális szubjektum</i> szétoszja önmagát szubjektumra és objektumra			
Hegel	az ÉN az észben, a legteljesebb valóságot <i>abszolút szellemben</i> megszüntetve-megőrizve és felemelkedve van jelen	az egyéni, individuális tudat az abszolút szellem (Isten) öntudatra jutási folyamatainak a szerves része			
félvilágosodás, liberalizmus			értelem, személyi szabadság; ateizmus		

jellemzők						
	nincs önálló én, a személyiség a világszellem része	a világ szétvált objektumra és szubjektumra, a lélek önálló entitás	az ember áll a világmindenség középpontjában az önző ember	csak az én létezik, a világ általa teremtet illúzió	a tudat az egy terméke	az én nem egységes, komponensekből áll
vallások, filozófiák, ideológiák (időrendben)						
kapitalizmus				szollipszizmus		
Berkeley				„a világ az én képzetem”		
Schopenhauer				az egyén nincs alárendelve a történelemnek; mindenk előtt az egyes egyén konkrét létfeltételeit kell fel-fogni; azaz „magamat a létezésben mértem”		
Kierkegaard				az <i>individuum</i> az ember fölötti ember; az <i>Übermensch</i> , aki a hagyományos értékektől való rökéletes szabadságával tűnik ki; vele szemben állnak a csordá-emberek		
Nietzsche						
kommunizmus, szocializmus(2)						
egzisztencializmus (Sartre, Jaspers, Heidegger)				az egyéni lét megelőzi a lényeket		
természettudományos nézetek:						
- Freud pszichoanalízise					- felettes én (<i>superego</i>) - én (<i>ego</i>) - ösztönén (<i>id</i>)	
- Jung komplex pszichológiája(3)					<i>persona</i> – tudat – kollektív tudattalan	

	jellemzők				az <i>én</i> nem egységes, komponensekből áll
	nincs önálló <i>én</i> , a személyiség a világszellem része	a világ szétvált objektumra és szubjektumra, a lélek önálló entitás	az ember áll a világmindenség középpontjában	csak az <i>én</i> létezik, a világ általa teremtet illúzió	
vallások, filozófiák, ideológiák (időrendben)					
- emergens felfogás (Kurzweil, Harari)					x
- kontinuum felfogás (Kaku)					x
- modern neurobiológia (ide sorolható Kahneman is)					nincs egységes <i>én</i> , a személyiség folyamatos változásban van, melyet neurobiológiai folyamatok véletlen-szerű kilengései okoznak, döntéseink ösztönösek
posztmodern, dekonstrukció(4) (Foucault, Derrida)					az <i>én</i> fogalma relatív, tartalmazza a másik ént is, éniünk sokrétű
kiberfilozófia(5) (Wiener, Ashby, Kurzweil, Günther)					a szerves és a szervetlen szubjektum között fokozatos az átmenet

Megjegyzések:

- (1) A *Buddhizmust* ebbe a kategorizálásba nem tudjuk bevonni. Fő tanítása úgy hangzik, hogy csak a tudat létezik, a világ általa teremtet illúzió, beleértve az ÉN-t is. Személyiségünk létezése megalapozatlan feltételezés; a létforragnál csak úgy szabadulhatunk ki, ha megszabadulunk ettől az illúziótól.
- (2) A *kommunizmus* és a *szocializmus* a személyiséget mint osztályának tagját tekintti, e rendszerbe nem sorolható be.
- (3) Jung besorolása is kétséges; a kollektív tudattalant akár „a személyiség a világszellem része” felfogás inverzének is tekinthetjük.
- (4) A *posztmodern* és a *dekonstrukcionista* filozófia a személyiséget – mint mindent – relativizálja, összetevőire bontja, elveti a személyiség metafizikus felfogását.
- (5) A *kiberfilozófia* – félredobva az arisztotelészi duális logikát és ontológiát – a személyiséget szétbonthatóan tartja tetszés szerinti számú összetevőre, és az élettelen, a gépet, a nem biológiai entitást is felruházhatóan tartja olyan jellemzőkkel, melyeket eddig csak élő szervezeteknek tulajdonítottunk. Ezzel feloldja a határt élő és élettelen között. A kiberfilozófia rokon vonást mutat a posztmodernnel és a dekonstrukcióval, mely a személyiséget relativizálja és szinten feldarabolja.

Forrás: a táblázat Héjjas Irván: Távolságtól vallások és filozófiák c. tanulmánya (lásd e kötetben), valamint *Isten véled, materializmus!* (Budapest, Eötvös József Könyvkiadó, 2017.) és *A megismerés útjai* (Budapest, Bioenergetic Kiadó, 2009.) c. könyve felhasználásával, valamint a vele folytatott konzultáció alapján készült.

A személyiség a különböző vallásokban, filozófiákban és ideológiákban (időrendben és főbb jellemzők szerint – táblázatfüggvény)

					x	kiberfilozófia(5)
					x	posztmodern, dekonstrukció(4)
				xx		természettudományos nézetek
				x		- Jung (pszichológiája)(3)
				x		- Freud (pszichoanalízise)
			x			egzisztencializmus
						kommunizmus, szocializmus(2)
			x			Nietzsche
			x			Kierkegaard
			x			Schopenhauer
			x			Berkeley
		x				kapitalizmus
		x				felvilágosodás, liberalizmus
x	x					Hegel
x	x					klasszikus német idealizmus
	x					reformáció, protestantizmus
	x					muzulmán vallás
	x					kereszténység
	x					Arisztotelész
x	x					Arisztotelész előtti görög filozófiák
						buddhizmus(1)
		x				számkhya bölcelet (hindu)
		x				upanisadok (hindu)
x	x					védánta bölcelet (hindu)
x						konfuciánus filozófia
x						taoizmus (a Wu tudás)
	x					zsidó vallás
nincs önálló én, a személyiség a világ-szellem része	a világ szétvált objektumra és szubjektumra, a lélek önálló entitás	az ember áll a világ-mindenség közép-pontjában	csak az én létezik, a világ általa teremtett illúzió	a tudat az agy terméke	az én nem egységes, komponensekből áll	vallások, filozófiák, ideológiák (időrendben) főbb jellemzők

Megjegyzések a táblázatfüggvényhez

A két „tengelynek” valójában fordítva kellene lennie; tehát az időnek a vízszintesen (független változó), és ahogy haladunk előre az időben, úgy erősödik az újabb vallásokban, filozófiákban és ideológiákban az én, a személyiség eleme (függő változó). A „tengelyeket” technikai okok miatt cseréltem fel: nehéz lett volna a nagyszámú vallást, filozófiát és ideológiát a vízszintesen elhelyezni.

A buddhizmus, a kommunizmus és szocializmus nem sorolható be ebbe a kategorizálásba; Jungé is kétséges. (Lásd az előző táblázat alatt részletesebben.)

Facit

Kiinduló hipotézisünk szerint a nyugati civilizáció kiemelkedésében a személyiség kialakulása alapvető szerepet játszott, mely részben a tulajdonnal függ össze, de elsősorban a nyugati vallások, filozófiák és ideológiák talaján formálódott ki, és a folyamat csúcsa a személyiségi jogok széles körű biztosítása volt a nyugati polgárok részére.

Ennek megfelelően ha e folyamatot olyan függvény segítségével ábrázoljuk, melynek függőleges koordinátája mentén a különféle vallásokat, filozófiákat és ideológiákat tüntetjük fel kialakulásuk idejének sorrendjében, a vízszintesen pedig azokat a főbb jellemzőket, melyek az individualizáció felé mutatnak, akkor ahogy az időben egyre feljebb jutunk a modern korok felé, a „görbének” emelkednie kell.

Ezt ábrázoltuk „táblázatfüggvényünkben”, mely az előző, részletes táblázat összegzése. A trend alapvetően megfelel az elvárásainknak, de lényeges törések vannak rajta. Egyrészt a legősibb, a zsidó vallás (a szaduceusokat kivéve) már az egyéni lélek létezését vallotta (mely alapfeltétele az én fogalmának). De a törést főképpen az okozza, hogy az indiai vallások és bölcseletek egyes változatai – mint láttuk a részletes összevetéseknél – már a korai időkben is elismerték az én valamilyen formában való külön létezését. (Nevezetesen: a *védánta bölcselet*, az *upanisadok tanításai és a számkhja bölcselet*.)

A protestantizmus a kiválasztottsággal, a hivatástudattal, a munkán keresztül való üdvözüléssel igen erőteljes lökést adott az európai személyiségfejlődésnek. (Szemben a katolicizmussal, mely a szemlélődést és az önsanyargatást vallotta, bár a szabad akarat dogmájának hozzájárulását az individualizációhoz nem lehet túlbecsülni.) A felvilágosodás pedig az értelem uralmához, az emberi szabadság fontosságához, az ember középpontba állításához, istentagadó filozófiákhoz és az egzisztencializmushoz vezetett el. De innen ágaznak el a kommunizmus, a szocializmus és a kapitalizmus eszméi is.

Ugyanakkor – meglepő módon – a klasszikus német ideológia (Fichte, Schelling, Hegel) felfogása az abszolút szellemről kategorizálásunk szerint, mely a személyiség fogalmának kialakulását állítja középpontba, visszalépést jelent a távol-keleti vallások és filozófiák felé.

Nehezen illeszthető be kategorizálásunkba a buddhizmus, hiszen miközben az *én*, a *személy* létét tagadja, ugyanakkor azt vallja, hogy csak a tudat létezik, a világ a tudat által teremtett illúzió, beleértve az *ént* is.

A kommunizmus és a szocializmus a személy osztályhoz kötöttségét hangsúlyozta, nem mint egyéniségről beszélt róla. A klasszikus nyugati polgár – mint fogalom és mint életmód – a múlt század legelején érte el csúcspontját. (Lásd Márai kassai polgárát.) A század második felében a fogyasztói társadalomban már manipulált, elfajzott változatával állunk szemben.

A posztmodern, a dekonstrukció és a kiberfilozófia személyiségfelfogása valójában hanyatlást jelent az előzőekhez képest. Fejlettebbnek ezeket az ideológiákat

csak akkor tarthatjuk, ha olyan jövővel számolunk, amikor az emberi személyiség és a gépi intelligencia összeolvad, és az értékelésünk szerint előrehaladást jelent.



A nyugati civilizáció kiemelkedése alatt közkeletűen *fejlődést* értünk. Ez a fejlődés azonban mára gyógyíthatatlan sebeket ütött a globális ökoszisztémán, miközben a hagyományos, „nem fejlődő” társadalmak megkímélték azt. A nyugati fejlődés az életet sok vonatkozásban jobbra és komfortosabbá tette, az éhezést és a járványokat felszámolta, az átlagos életkort nagyban meghosszabbította, az állandó háborúskodás állapotát megszüntette, de kiélezte a társadalmi különbségeket és felborította a harmóniát, mely a hagyományos közösségekben az együttélést és a természethez való viszonyt jellemezte. Az *én*, a *személy* fogalmának kialakulása óhatatlanul az önérdek, majd az önzés előtérbe állításával járt, és ennek a fejleménynek a kapitalizmus rendszere felelt meg a legjobban. A nyugati kapitalizmus oly vonzó alternatívát kínált, hogy annak a hagyományos társadalmak sem tudtak ellenállni; ősi bölcseségeik és vallásaik sorban megbuktak az új „vallással” szemben. Manapság már ők járnak élen a „fejlődésben”, azaz a természeti környezet pusztításában és az egyén elmagányosodásában.

Robotok legyünk, vagy szabadok?

Nem a múltba révedés tesz minket bölccsé, hanem a felelősség a jövő iránt.

(G. B. Shaw)

A mesterséges intelligencia hasznos eszköz vagy végzetes veszélyforrás? Elősegíti önmegvalósításunkat, vagy emberi mivoltunkat fenyegeti? Hol kell meghúzni a határokat, hogy az információtechnológia a humánusot szolgálja? Az új korszak szempontjából deficitese számítógépek vagyunk, vagy továbbra is kivételes lények maradunk, akik értelmet, szépséget és érzelmet visznek az univerzumba? Ezek korunk és az elkövetkező évtizedek legfontosabb kérdései, a klímaváltozás mellett.

A személyiség jövője az evolucionista világmodellben

Ray Kurzweil szerint az univerzum evolúciója az információ testet öltése egyre komplexebb formákban.¹ Az első korszak a fizikáé és a kémiáé, amikor az információ megjelenik az atomi struktúrákban. A második a biológiáé, amikor az információ megjelenik a DNS-ben. Ezután az információ az idegsejtmintázatok formáját ölti (azaz kialakul az agy). A negyedik a „technológia” korszaka, a hardver és a szoftver megjelenése. És ekkor következik be az a korszak, a szingularitás, amely Kurzweil szerint előttünk áll: „A technológia és az emberi intelligencia egybeolvad: a biológiai módszerek (köztük az emberi intelligencia) integrálódnak az (exponenciálisan) bővülő emberi technológiai bázisba.” A hatodik a kiteljesedés korszaka: a világegyetem öntudatra ébred, „a világegyetemben létező anyag és energia átitatódik intelligens folyamatokkal és tudással”.² Vulgarizálva: Kurzweil információja Hegel istene; amikor az abszolút szellem – létrehozva a természetet, majd a különféle tudati szférákat – önmagára ismer. De ez a felfogás a buddhizmussal is rokonságot mutat, mely szerint a világ nem más, mint univerzális tudat. Hasonló végcélt mutat fel Teilhard de Chardin is: végső soron minden – tudás, energia, anyag – isteni egyesülés során eggyé válik – ez lesz az „Omega pont”. Kurzweil szingularitása még azt megelőző folyamat: amikor a mesterséges intelligencia utoléri, majd leahagyja a természetes intelligenciát; ember és technika egyesül, és attól kezdve minden megváltozik.

¹ Kurzweil információ alatt mintázatokat ért, intelligencián a mintázatok felismerésének képességét.

² Ray Kurzweil: *A szingularitás küszöbén*. 23.

Előttünk áll tehát a nagy kérdés: ez lesz a jövőnk? Összeolvadunk az általunk létrehozott technológiával? Mennyire törvényszerű ez a jövő? Kb. hetvenezer évvel ezelőtt a *Homo sapiens* képessé vált az elvont gondolkodásra.³ Addig a biológiai evolúció igen lassú, döcögős útján haladtunk. Ezzel ráfordultunk a kulturális evolúció gyors sztrádájára. Most pedig a technológiai evolúció szédületes gyorsaságú pályáján száguldunk tovább? Jó lesz ez nekünk? Akarjuk ezt a jövőt? Egyáltalán, módunkban áll nem akarni?

A szubjektum relativizálása

A kiberfilozófia ennek a jövőnek a megvalósításán dolgozik. A kibernetika és a digitalizáció nagy kihívást jelent a személyiség hagyományos értelmezésével szemben. A kiberfilozófusok ugyanis – Gotthard Güntherrel és Ray Kurzweillel az élen – a személyiség fogalmát ki akarják terjeszteni a gépre, nem biológiai alapú személyiségről és tudatról beszélnek. Nick Bostrom egyenesen rasszizmussal vádolja azt, aki nem akarja elfogadni a nem szénalapú élet lehetőségét. E szándéknak reális alapja van. A leginkább figyelemre méltó érvekkel talán Max Tegmark szolgál. Azt állítja, hogy intelligencia és nem intelligencia között nem lehet éles határt vonni. Továbbá: bebizonyítja, hogy „az információ a fizikai közegtől függetlenül saját életre kelhet”.⁴ A számítás, a memória és a tanulás „közegfüggetlen” funkció, szervesen anyagokkal is végrehajtható. Ebből az következik, hogy „a MI lehetséges: az intelligenciának nincs szüksége hús-vér testre vagy szénatomokra”.⁵

Günther a személyiség gépi kiterjesztését nem tradicionális, nem arisztotelészi, többértékű logika és ontológia kifejlesztésével alapozza meg. A szubjektumot szétválasztja nem felosztott és felosztott változatra. A szubjektum első formája az egyén, a második „az általános szubjektivitás interszubjektív médiuma”. Ezzel a szubjektum, a szubjektív fogalmát demisztifikálja, deszakralizálja és természet-tudományos alapokra helyezi, és olyan elméletet dolgoz ki – ez a kibernetika tárgyköre –, amely a szubjektív tulajdonságokat (öntudat, önreflektáló képesség, gondolkodás, döntéshozatal stb.) mesterségesen is képes előállítani.⁶

A szó szoros értelmében az orrunk előtt játszódik le a *Homo sapiens* eljelentéktelenítése. A robotok először elveszik a fizikai munkánkat, majd a mesterséges intelligencia kiszorítja az embert a szellemi tevékenységekből is, végül azáltal, hogy a gépeknek is személyiségük lesz, és kifejlesztik a nem biológiai alapú tudatot is, az ember megszűnik a világ unikális teremtménye lenni, egyenlő lesz a gépekkel. (Melyeket viszont ő alkotott! És melyek Stiegler szerint emberi lényegünk!)

³ Ez Harari becslése. (*Sapiens*.)

⁴ Max Tegmark: *Élet 3.0 – Embernek lenni a mesterséges intelligencia korában*. 68.

⁵ Uo. 77.

⁶ Lásd erről Günther műveinek ismertetését a Mellékletben.

A személyiség jövője a világról alkotott más gondolati konstrukciókban⁷

Az emberiség legősibb gondolati konstrukciója az Egészben a Rész, a Részben az Egész típusú gondolkodás volt. Az egyistenhívő vallások az Egészben a Rész jellegű modellt testesítik meg. Mai világunkra, a felvilágosodás óta, a materializmus és az evolúciós gondolkodás eredményeképpen a Részekből felépülő Egész logikája jellemző. Ez a gondolkodás – mint fentebb láttuk – törvényszerűen vezet el a világmindenség evolúciójához. Azt követően, hogy Arisztotelész kivette az egyént az egészből, szembeállította vele, s kétértékű logikája megkülönböztette az ént a minden mástól, teret adott annak a fejlődési folyamatnak, amelyben a gépek, a technológia önállósodik, és az emberrel szemben álló tényezővé válik. A számítógép saját lábra áll, a szoftver önmagát fejleszti, tökéletesíti, kinő a programozó kontrollja alól. Ahhoz, hogy ne így legyen, hogy ez ne történjék meg, a világról való gondolkodási módunkat kellene gyökeresen megváltoztatni... Ami lehetetlennek tűnik.

Ember és gép

Közkeletű meghatározás szerint a mechanikus gépet az különbözteti meg az élő szervezettől, hogy komponenseiből összerakható és működése determinált. Ezzel szemben az élő szervezet nem rakható össze részekből, kifejlődik, „kinöveszti” önmagát. Működése, jellegzetességei nem előre meghatározottak (a DNS-ben csak a fontosabb fizikai jellemzők és a hajlamok vannak programozva).

A mesterséges intelligencia ma már semmiképpen sem nevezhető hagyományos értelemben vett gépnek; elindult az élő szervezetekre jellemző tulajdonságok kifejllesztése útján. A *deep learning*, a gépi tanulás ennek jeleit viseli: a program önmagát fejleszti, az adott adatbázis feldolgozási igényeihez igazodva. És nem is determinisztikus, hiszen rendelkezik bizonyos gondolati, döntési önállósággal; és annak függvényében más-más eredményeket fog „kiadni”.⁸ A go ázsiai nagymesterét, világbajnokát legyőző AlphaGo önmagától tanulta meg a játék szabályait: megmutattak neki bizonyos számú játszmát, melyből leszűrte, mik a szabályok.

A metafizika és a vallásos gondolkodás talaján állók arra hivatkoznak, hogy az élettelen anyagban nincs meg az a valami, ami az élőtlé élővé teszi, amit a Teremtő vagy a világszellem az élettelen anyagba „belerak”. És valóban, élő szervezetet

⁷ Lásd Sütő Zoltán két tanulmányát e kötetben: Gép és ember viszonyának szemléleti kérdései (a II. fejezetben) és A mesterséges intelligencia lehetőségeinek értelmezése különböző világmodellekben (a IV. fejezetben).

⁸ Ha a szoftvert nagyobb adatbázison trenírozzák, mások lesznek az eredmények. Itt félelmetes párhuzam állítható fel: attól függően, hogy az embert kis korában milyen hatások érik, és milyen társadalmi közegben szocializálódik, más ember lesz belőle felnőttkorában, függetlenül attól, hogy milyen biológiai „csomagot” hozott magával. (Ez utóbbi csak a hajlamait határozza meg.)

művi úton mindeddig nem sikerült létrehozni (bár Craig Venter 2010-es kísérlete már nagyon közel jár ehhez).⁹ Sütő Zoltán ezt a valamit belső hajtóerőnek (hajtás, űzés, késztetés, angolul *drive*) nevezi, és szintén metafizikai eredetűnek tartja.

Ha az élet, életösztön, mozgó szellem eredetét isteninek és metafizikainak tekintjük, akkor tovább bonyolítjuk a helyzetet: az evolúció minden nagyobb ugrásánál feltételeznünk kell ezt a külső beavatkozást (így a többsejtű élőlények, majd a helyváltoztató állatok kialakulásánál, végül a tudat és az öntudat létrejöttéért). Evolúciós alapon viszont ezek a nagyobb változások „logikusan” történnek meg, és az anyag emergens tulajdonságával magyarázhatók.¹⁰ Kaku „kontinuumelmélete” meggyőzően bizonyítja, hogy az érzékelési-visszacsatolási hurkok számának és komplexitásának növekedésével hogyan lesz egyre komplexebb az élővilág, és alakul ki végül a tudat, majd az öntudat.¹¹

E kontinuitáselmélet alapján nincs okunk azt feltételezni, hogy ez az ugrás az emberi agy és a gépi intelligencia között ne történhetne meg, hogy az agy alapján nem lenne kifejleszhető az emberi agyat utólró, majd túlszárnyaló mesterséges intelligencia. Gotthard Günther kiberfilozófus munkásságának középpontjában ez a törekvés állt. Az identitást „saját jogon fellépő technikai problémának tekintette”, és lehetségesnek tartotta a tárgyak identitásának meghatározását is.¹²

Gondolkodik-e a gép?

2018-ban a Tesla önvezető autója próba közben halálra gázolt egy kerékpárját a zebrán áttoló nőt Arizonában. Hogy történhetett ez meg? Az autó mesterséges intelligenciája nyilván felismeri az embert, meg a kerékpárt is, meg a kerékpárját toló embert is, meg azt is tudja, hogy a zebrán a gyalogosnak van előnye. De ezek szerint hiányzott az adatbázisából az a „kép”, amikor a gyalogos a zebrán tolja a biciklijét. Ez a szituáció számára ismeretlen volt. A programozás újabb kapitális hibája: akkor miért nem állt meg? Ezek szerint ha egy tevehajcsár megy át a zebrán a tevéjével, azt is elüti. (Mert Arizonában ez nem tartozik a lehetséges forgatókönyvek közé.) Ez az algoritmus nem gondolkodik, csak „mintafelismerésre” képes; és csak olyan mintákat ismer fel, melyeket adatbázisába betápláltak. Előfordult, hogy azért történt súlyos légi katasztrófa, mert a robotpilótába az adott szituáció nem volt betáplálva.

Watson, az IBM szuperintelligens számítógépe sem gondolkodik, amikor kórházi alkalmazása során jobb diagnózist állít fel, és jobb terápiát javasol, mint egy orvosi konzílium. De elképesztően nagy memóriájába olyan rengeteg vonatkozó esetet betápláltak, amennyi felülmúlja az orvoscsapat ismereteit. Kiszűri, hogy milyen tüneteket milyen betegségként azonosítottak, és melyik kezelési

⁹ Venter mesterségesen létrehozta egy baktérium DNS-ét, majd azt beültette egy másik fajta baktériumba; az eredmény: a baktérium tovább élt a mesterséges DNS-sel – új faj keletkezett.

¹⁰ Lásd ennek részletes kifejtését Kiss Károly A tudat rejtélye c. tanulmányában, a III. fejezetben.

¹¹ Lásd uo.

¹² Lásd erről a tanulmányaiból közzétett ismertetést a Mellékletben.

gyógymód volt a legcélravezetőbb. A fentebb említett AlphaGo esetét tekintve a számítástechnikai szakma szerint az algoritmus nem gondolkodott, hanem összefüggéseket, korrelációkat keresett, és ezek alapján „találta ki” magától a játék szabályait. (De mi ez, ha nem gondolkodás? – kérdezheti a laikus ember – magamat is beleértve.)

Megismerhető-e az agy?

Szoftveroldalon az emberi agy (de szinte bármilyen agy, még a rovarok kezdetleges agya) működésének algoritmikus leírása is a vártnál sokkal nehezebb problémát okoz. Évekbe telt, amíg a *C. elegans* nevű fonalféreg 302 idegsejtéből álló „agyának” a működését pontosan le tudták írni – holott neki csak 7000 szinapszisa van. (Szemben az emberi agy 80–100 milliárd neuronjával, melyek között 10^{14} , 10^{15} – azaz százezer milliárd vagy akár millió milliárd – szinapszis alakulhat ki.) Ráadásul az agy más elvek szerint működik, mint a számítógép. Az agy számítási módja analóg, a számítógépé bináris. A számítógép szekvenciálisan működik (azaz minden lehetséges változatot „befut”), az agy – ha megvan a megoldás – a maradék irányokra nem pazarol energiát. A neokortexben a mintafelismerést végző neuronok párhuzamosan dolgoznak. Ha az agy tényleges működését szimuláló számítógépet akarnának építeni, annak mérete világváros nagyságú lenne, és hűtéséhez egy nagy folyó vizére lenne szükség.¹³

Gothard Günther szerint egy bonyolult rendszer csak fölötté álló szintről érthető meg.¹⁴ A világ felépítése hierarchikus, állítja Kurzweil, minden be van ágyazva a fölötté lévő magasabb szintbe. De vajon mi az a magasabb szint, melybe az emberi agy be van ágyazva, és ahonnan az ismerhető? Heisenberg *A rész és az egészben* kifejti, hogy szerinte téves az a felfogás, hogy minden ismerhető, és ha valamit még nem értünk, az azért van, mert a tudomány még nem tart ott, az csupán idő kérdése. Példának a rádium-B atomot hozza fel: nem lehet előre megmondani, hogy mikor bomlik fel; mikor lök ki magából egy elektront, és milyen irányban. Szerinte a természet itt megálljt parancsol – a további részletek és okok nem ismerhetők meg.

Identitás

Az informatikusok és MI-szakértők többsége szerint a gép minden olyan tulajdonsággal, jellegzetességgel „felprogramozható”, melyeket eddig az emberi személyiség sajátosságának tekintettünk. Ezzel szemben Sütő Zoltán azt állítja, hogy

¹³ Az agy és a számítógép működésének összevetéséről lásd Kiss Károly *A tudat rejtélye* c. áttekintését a III. fejezetben.

¹⁴ Günther munkáinak összefoglaló ismertetése a Mellékletben olvasható.

a tárgyaknak, így a gépnek, a MI-nak nem lehet identitása, azaz személyisége. Identitás alatt a következőt érti:

„A tudomány nem véletlenül nem kísérletez az emberi identitás modellezésével, e téren csak a művészetek képesek eredményeket felmutatni. Kiemelkedő képzőművészeti alkotások (lásd portréfestészet) vagy jelentős szépirodalmi művek tárgyaként megjelenítve tudhatunk meg lényeges jellemzőket egy személy identitásáról, mely jellemzők mindenkor a személy drámájában nyilvánulnak meg. A személy drámájának két pólusa, mégpedig a »boldogság/öröm« és a »tragédia/szenvedés« között válik láthatóvá a személyből az, ami állandó, vagyis az identitás, mivel boldogság azért érheti a személyt, mert közelébe kerül annak, ami belső lényegében pozitív tartalom, illetve tragédia éppen azért éri a személyt, mert elveszti azt, ami belső lényegében pozitív. Az identitás állandósága éppen abban érhető tetten, hogy pl. a tragédiában nincs mód a személy pozitív belső tartalmát korrigálni a veszteség elviselhetőbbé tétele érdekében, sőt, a tragédia éppen azért marad elháríthatatlanul tragédia, mert a személy pozitív belső dinamikája, vagyis identitása változhatatlan, és a tragédiában pótolhatatlan veszteség érte.»

„[...] az ember éppen ettől ember. Ha valamilyen beavatkozással megszüntetnénk identitásának állandóságát (akár a tragédiák kiküszöbölése végett), lényegileg szűnne meg az így kapott »konstrukció« embernek lenni. Ettől válna gépezetté. Továbbá a három absztrakt világmodellből a középsőben, azaz a Rész(ek)ből épülő Egész esetében egy ilyen komplex identitás értelmezése, mely minden személy esetében teljességgel egyedi, nem tűnik lehetségesnek. Az így felfogott identitás nem értelmezhető az evolúció logikájában. Az evolúció szempontjából nem tűnik egyébnek a felvethetősége, mint egy túlélőgépezetnek. Az evolúcióban nincs ok az egyedet drámával terhelni [...]»¹⁵

„A másik két világmodellről viszont nem idegen az identitás ilyen felfogása, lévén az egyikben az identitás teremtési adomány, a másikban a Teljesség artikulációja a belső *készítésben*.»

„[...] Meglehetősen kétségesnek tűnik, hogy az ember modellezésére vállalkozó MI tisztában lenne a feladat komplexitásával. Az informatikához közelebb eső nyelvezetre lefordítva a MI embermodelljének rendelkeznie kell olyan sajátos és egyedi vezérléssel (programmal), ami nem véletlenszerű, ugyanakkor saját vonatkozásában határozott létezési koncepció, mely koncepciónak semmilyen más reprezentációja nem létezik, vagyis nem származtatható más egyedből. Kétségeink vannak aziránt, hogy ilyen tulajdonságokkal rendelkező humanoid megalkotására a mai értelemben vett MI alkalmas lehetne. Továbbá nem gondolnánk, hogy a transzhumanista álom valaha is tartalmazhatna ilyen komplexitású megvalósítási igényt. Ezzel szemben meglehetősen nyilvánvalónak tűnik, hogy a transzhumanista mennyországban a gépekkel összeépülő ember, »aki«

¹⁵ Az evolúció elméletéből nem ez következik. Ez az okfejtés minden olyan esetre vonatkozik, amikor öröm, fájdalom vagy bármilyen érzés tudatosul. A tudatosulás feleslegesnek látszik, hiszen a szervezet vegetatív (azaz öntudatlan módon) mindenképpen reagál ezekre a hatásokra. Ezen érzetek tudatosulásának csak az lehet az evolucionista magyarázata, hogy *van evolúciós funkciójuk*. (Kiss Károly szerkesztői megjegyzése.)

minden felmerülő vágya elől technikailag hártja el az akadályokat, éppen a drámát, tehát valóságos emberi mivoltát veszti el.”¹⁶

Így a MI lehet intelligens, rendelkezhet több, a szubjektumra jellemző tulajdonsággal, de identitása nem lehet – állítja Sütő Zoltán. Ezen a véleményen vannak a szoftverfejlesztő mérnökök is. Belátható időn belül képesek leszünk kifejleszteni az emberit meghaladó szuperintelligenciát, de annak nem lesz identitása, nem lesz tudata, nem lesz „személy”.

Telek, nyarak, mintafelismerés, adatbázis¹⁷

A mesterséges intelligencia fejlesztése során adódnak hol optimizmusra okot adó nagy ugrások, hol elakadások („telek”), borúlátó szakaszok. A Moore-törvény pl., mely azt mondta ki, hogy a számítógépek kapacitása másfél-két évenként megduplázódik, sokáig érvényben volt, de 2015 óta ez a növekedés lelassult, a szilíciumalapú technológia kifulladt, és valószínűleg teljesen új technológia fogja csak megint fellendíteni a számítási kapacitás növekedését.¹⁸

A szoftverek fejlesztésében az utóbbi években a mesterséges neuronhálózaton alapuló „gépi tanulás” a legelterjedtebb módszer. Ennek lényege a mintafelismerés, az osztályozás és az összefüggések megállapítása. Mindhárom esetben nagy adatbázisra van szükség, melyből az algoritmus tanul, melyen „trenírozzák”. A mintafelismerés úgy történik, hogy az algoritmus statisztikai, valószínűség-számítási módszerekkel beazonosítja, hogy a vizsgált kép, jelenség, szó, szöveg, összefüggés az adatbázis melyik mintájára hasonlít a leginkább. A gépi tanulás elterjedésének legnagyobb akadálya most a szűkös adatbázis. Ez állhat természetes és virtuális (mesterségesen kreált) adatokból (képekből, mintázatokból). Az Amazon emberi személyzet nélküli bolthálózatot működtet. Eddig 26 „Go” üzletet nyitott meg. E hálózat működtetéséhez olyan adatbázisra van szükség, amely a szenzorok által a vásárlók minden potenciális mozgását, gesztusát be tudja azonosítani. (Az áru levétele a polcra, megsemmisítése, megforgatása, kosárba rakása, esetleg visszatevése, esetleg nem a kosárba tévése, hanem elrejtése stb.) Léteznek a MI alkalmazásához szükséges adatokat tömegesen gyártó vállalatok is. A kínai MBH pl. 300 ezer (!) alkalmazottal dolgozik, akik az élet legkülönbözőbb területeiről származó, szinte végtelen mennyiségű képet „felcímkézik”, hogy a gépi keresés, osztályozás eredménye minél pontosabb legyen.

A szoftverfejlesztésnek ez az iránya nagyon megnöveli a hardver energiafogyasztását. A Google 2018-ban épített BERT nevű nyelvi mesterséges intelligenciája adatbázisként a Wikipedia 3,3 milliárd nagyságú szókészletét használja fel.

¹⁶ Az idézetek Sütő Zoltán A mesterséges intelligencia lehetőségeinek értelmezése különböző világmodellekben (a IV. fejezetben) c. tanulmányából származnak.

¹⁷ Az *Economist* 2020. június 13-i technológiai melléklete alapján. „Steeper than expected.” „Artificial intelligence and its limits”.

¹⁸ Lásd erről Kiss Károly: *Ki akar itt Superman lenni?* (Az Információtechnológia c. pont.)

Egy-egy „tréning” több millió dolláros villanyszámlába kerül.¹⁹ A kvantumkomputerek működését a szakemberek elméletileg már megértették, megoldották, de a megépítés súlyos mérnöki problémákba ütközik.

A PWC felmérése szerint 2020-ban a vállalatvezetők 4 százaléka tervezte, hogy MI-t fog alkalmazni – 2019-ben ez az adat még 20 százalék volt. Az azt ténylegesen alkalmazók aránya 27-ről 18 százalékra esett. A felmérés során megállapították, hogy „eltűnőben van az az irreális várakozás, mely az elmúlt éveket jellemezte”. A MI alkalmazása nem minden területen eredményes. „A gépi tanuláson alapuló módszer könnyedén meg tudja mondani, hogy a go játékban adott táblaállásnál bizonyos lépésnek hány százalék a győzelmi esélye. Ugyanakkor nehezen tud kezelni olyan helyzeteket, amikor a repülőtéri bevészolgalatnál *chat-boxként* alkalmazva egy utas panaszával kell foglalkozni, akinek elmaradt a járata” – idéz az *Economist* egy ismert MI-szakértőt.²⁰

Alávetettség²¹

Széles körben elterjedt a félelem, hogy a mesterséges intelligencia uralma alá fogunk kerülni. Először talán Martin Rees pendítette meg ezt a lehetőséget,²² de később egy sor ismert tudós (Hawking, Musk, Bostrom) is tett ilyen kijelentést. Ha a mesterséges intelligenciának kialakulna a tudata, akkor ez a fejlemény könnyen elképzelhető lenne. De nagyobb a valószínűsége annak, hogy emberek egy csoportja, vagy még inkább egy ország a szuperintelligencia felhasználásával hódítja meg a világot, veti uralma alá a többi embert. Hogy ez hogyan következhet be, arról Max Tegmark ír igen hihetőnek tűnő forgatókönyvet *Élet 3.0* c. könyvében.

Közelebbi és reálisabb az a félelem, hogy a MI elveszi a munkánkat. Sok fizikai munkát és szellemi rutinmunkát már ma is gépek, algoritmusok végeznek, a gépi intelligenciával rendelkező robotok térhódítása nyilvánvaló.²³ Nem nehéz ebből arra következtetni, hogy lassan-lassan feleslegessé válunk. Márpedig emberi tartást, önbecsülést a munka ad, és ha ez nem is okozna egzisztenciális-anyagi problémát (az élőmunka helyett a robotokat adóztatnák meg, és aktuálisná válna a már ma sem politikai fantazmagóriának tűnő minimális, állampolgári jogon járó jövedelem), a folyamatot tömeges lelki megbetegedések kísérnék. Ekkor időszerűvé válna Huxley jóslata a *Szép új világban*, melyet ő a XXVI.

¹⁹ *The Economist*, 2020. június 13.

²⁰ A számítástechnikában ezt Moravec-paradoxonnak nevezik: a gép könnyen megold összetett matematikai és logikai feladatokat, de a mindennapi élet rutinjelenségeit nehezen tudja értelmezni és utánozni.

²¹ Csak hogy más jellegű asszociációkat is felkavarjunk, az alávetettség arab fordításban: iszlám.

²² Rees neves angol matematikus és asztrológus magyarul is megjelent, *Ütött az utolsó órák?* c. könyvében fél tucat lehetőséget felsorol arra vonatkozóan, hogy hogyan irthatja ki magát az emberiség a saját maga által létrehozott technikával a XXI. században.

²³ Lásd erről Kiss Károly: *Ki akar itt Superman lenni?* (A Semmittevésre kárhoztatva c. pont.)

századra jövendölt meg. A feladat úgy szólna, hogyan lehet helyettesíteni hormonokkal, hangulatjavítókkal és más gyógyszerekkel vagy biológiai eljárásokkal azt az örömet és megelégedettséget, a hasznosság tudatát, melyet az értelmes munka vált ki. És hogyan lehet elkerülni a tömeges káros addikciót, mely ilyenkor törvénytörően jelentkezik.

Szuperintelligencia és boldog jövő

A félelemmel, a szkepticizmussal, a borúlátással szemben jelen van az öröm és az optimizmus is.²⁴ Lubickolunk az információbőségben, élvezettel böngésszük a netet, tobzódunk a kommunikációs lehetőségekben, szinte ki sem kerül a kezünk közül az okostelefon. Látjuk, hogy a nethasználat függőséget is okozhat, hogy gyermekeink milyen veszélyeknek vannak kitéve, de ezt bagatellizáljuk: az új kommunikációs eszközök megjelenése (rádió, tv) korábban is ezzel járt.

A MI hallatlan előnyökkel jár a mindennapi életben és szinte minden szakterületen. Az optimisták ezeknek az előnyöknek a fokozódását látják, és nem számolnak súlyosabb, alapvető következményekkel, különösen olyanokkal, melyek lényegi mivoltunkat, az emberi személyiséget veszélyeztetik. A folyamatot kezelhetőnek tartják, nem félnék attól, hogy a technológia irányítása kicsúszik a kezünk közül. Mindazt, amit az informatika, a mesterséges intelligencia világa nyújt, beilleszthetőnek tartják abba a folyamatba, mely az ipari forradalommal kezdődött el, mely jobbá teszi a világot, életünket. Véleményük szerint a MI e fokozatos műszaki fejlődésnek egy állomása. Nem érzékelnek törést abban, ha a mesterséges intelligencia eléri, vagy meghaladja az emberit. Úgy gondolják, hogy a gépi intelligencia nem veszi el a munkánkat, hanem elvégzi annak gépies, mechanikus jellegét, és nekünk több időnk marad az érdemi, kreatív tevékenységre. A műszaki fejlődés fő irányait az ember fogja meghatározni, a gépi intelligencia pedig inkább az új technológiák kidolgozását, megvalósítását, majd elterjesztését, mindenki által elérhetővé válását fogja elősegíteni. A szuperintelligencia nem a műszaki fejlődés csúcsa lesz, és nem a legfontosabb technológia. A MI változatlanul eszköz marad a kezünkben, még annak a szuperintelligenciát megvalósító változata is.

Sundar Pichai, a Google elnöke úgy fogalmazott, hogy a MI nagyobb hatású technológia, mint annak idején a tűz, majd később a villanyáram felfedezése volt. A PWC nevű nemzetközi piackutató cég számításai szerint 2030-ra a MI 16 billió (ezer milliárd) dollár nagyságú új értéket fog hozzáadni a világgazdaság teljesítményéhez – ez olyan lesz, mintha újabb, Kína nagyságú gazdaság keletkezett volna.²⁵

²⁴ Kötetünkben – amint azt az Előszóban is jeleztük – nem törekszünk a különböző nézőpontok összehangolására; célunk inkább ezek sokféleségének bemutatása.

²⁵ Az *Economist* 2020. június 13-i technológiai melléklete.

Analóg és bináris világok; a józan ész halála

Amennyiben az internet keresőmotorjait is a mesterséges intelligencia körébe soroljuk, jól ismertek annak manipulatív hatásai: a kereső személy érdeklődésének kiszolgálásával mintegy információs buborékba zárja őt. Választások idején számtalan lehetőség nyílik a szavazók befolyásolására a közösségi médiumokon keresztül. „A politikai kérdésekből egyre inkább a komputerek által kiszámolt és megválaszolt szakmai kérdések lesznek” – írja Precht.²⁶ E szociáltechnikák egyre inkább kiszorítják a laikus választót a politikai ügyek intézéséből, a liberális demokrácia fokozatosan a komputertudósok által irányított „expertokráciává” válik.

A MI munkaerőpiaci ígérete úgy szól, hogy bár számos munkakörben felváltja az emberi munkát, a feleslegessé váló emberek valójában magasabb szintű, több kreativitást igénylő és több hozzáadott értéket termelő tevékenységet fognak folytatni. Azonban egyre több az olyan eset, amikor nem az ember irányítja a komputert, hanem a technológia az embert. És – bár lesz rá lehetőség – nem lesz képes mindenki a gépeknél magasabb szintű, kreatív tevékenységet folytatni.

A fogalomalkotást, a jelenségek magyarázatát átengedjük a mesterséges intelligenciáknak. Világképünket már nem mi hozzuk létre, hanem a komputerek logikája. A gépi intelligencia digitálisan modellezi a világot. A valódi világ lassan összecsúszik a gépek által előállított és modellezett világgal. Ez már más, nem emberi, hanem gépi világkép lesz, ahol a józan észnek (*common sense*, vagy németül *gesunden Menschenverstand*) már nem lesz helye. Emberi ésszel egyre kevésbé leszünk képesek megérteni ezt a virtuális, komputerlogika által előállított világot, melyet már nem mi, hanem a gépek irányítanak: nem közvetlenül, hanem a gépi logika által. A mindennapi ügyektől a sorskérdésekig a gépek (a gépek logikája) fognak dönteni. Ily módon uralkodik majd a gép az ember fölött.²⁷

A technika ördöge

Ha azt vizsgáljuk, hogy mit tehetünk a MI térhódítása ellen, megállíthatjuk-e, vagy lassíthatjuk-e a folyamatot, elképzeléseinknek megfelelően irányíthatjuk-e, akkor abba a végső kérdésbe botlunk, hogy mi viszi előre a történelmet. Mi az ember, az emberiség szerepe a történelemben? Formáljuk, vagy elszenvedjük azt? A válasz döbbenetes; úgy tűnik, hogy akár objektív idealisták vagyunk, istenben vagy az abszolút szellemben hiszünk, akár materialisták, történelmünk csupán megtörténik velünk. Még a marxi formációelmélet értelmében is a történelmet a termelőerők és a termelési viszonyok dialektikája viszi előre, értve az utóbbin a gazdálkodás tulajdon- és szervezeti viszonyait. Az ember mint a termelőerők része, munkavégző képességeivel vesz részt a folyamatban. (A termelőerőkre és a termelési viszonyokra mint alapra épül fel a kulturális és szellemi felépítmény, mely

²⁶ Richard David Precht: *Künstliche Intelligenz und Der Sinn des Lebens*. 226.

²⁷ Lásd erről Precht idézett művének Die smarte Matrix c. fejezetét.

vissza is hat arra.) Egy-egy kor vezéregyéniségének szerepéről Tolsztoj mondja ki a verdiktet: nem ő formálja a történelmet, hanem ő a kiválasztott, aki beteljesíteni hivatott a történelmi folyamatokat.²⁸

Egyre inkább úgy érzem, hogy az új technológiák okozta veszélyek tanulmányozása is hasonló végkövetkeztetéshez vezet: nem vagyunk képesek kezünkben tartani a sorsunkat, a történelem átgázol rajtunk, fölöttünk. A technológiai fejlődés formájában olyan determináció valósul meg, amelyen nem tudunk változtatni. A technológia öntörvényű, az irányítja a történelmet. „Előbb megépítjük a szerszámokat, aztán azok építenek meg minket” – idézi Kurzweil Marshall McLuhant *A szingularitás küszöbén* c. könyvében. Ha ezt elfogadjuk, ha ebbe belenyugszunk, akkor a további kutatás témája a kármentés és az alkalmazkodás kell legyen: Mit csináljunk, hogyan éljünk olyan jövőben, melyet egyre inkább a mesterséges intelligencia ural majd. Melyet ugyan mi hozunk létre, de nem áll módunkban azt nem létrehozni és egyre tökéletesíteni. (Sokszor idéztem már ezzel kapcsolatban Madáchot: „Az ember ezt ha egykor ellesi, vegykonyhájában szintén megteszi.”)

A technológiáról írt tanulmányom ezt a tehetetlenséget tükrözi. Súlyosítja a helyzetet, hogy Bernard Stiegler, akit ma a technofilozófia legnagyobb alakjának tartanak, egyenesen azt állítja – ez életművének lényege –, hogy a technológia nem eszköz, nem „külső” valami, hanem emberi lényegünk (!). Attól váltunk emberré. És a legrelevánsabb kérdés: mivel pótolható majd – és pótolható-e – az a munka és küzdelem, mely eddig értelmet adott az emberi életnek, ha az intelligens robotok és szoftverek mentesítenek majd alóla.

A kétkedőket az optimisták álláspontja nem nyugtatja meg – mert az a gépi intelligencia jóhiszemű felhasználását tételezi fel. De mi a garancia arra, hogy emberek egy csoportja, vagy akár egy ország a szuperintelligenciát nem mások alávetésére használja fel? Továbbá: ha a gépi intelligencia képes az önfejlődésre/önfejlesztésre, akkor mi a garancia arra, hogy önfejlődése nem vehet emberellenes irányt? Sütő Zoltán és jómagam e folyamat veszélyeit hangsúlyozzuk, a gőzgéptől a MI-ig tartó fejlődést nem homogén vonalnak tekintjük, a most kifejlődő gépi intelligenciák szerintünk már más megítélés alá tartoznak. A nemzetközi szakirodalomban szintén a fenntartás, a félelem és a kétkedés az uralkodó álláspont.

A futurista baktérium jóslata a jövőről Kr. e. kétmilliárdban

A legtöbben alaptalan fantazmagóriának tartják azokat a jövőképeket, melyekben a gépi intelligencia dominál. A hihetetlennek tűnő, beláthatatlan jövőbeli változások illusztrálására Kurzweil két baktérium Krisztus előtt kétmilliárdban folytatott beszélgetését idézi fel. (A beszélgetés a jövőt látó, futurista baktérium és fantáziátlan társa között folyik.)²⁹

²⁸ Lásd erről a *Háború és béke* II. kötetének bevezető filozófiai fejtegetését.

²⁹ Ray Kurzweil: *A szingularitás küszöbén*. 280–281.

Két milliárd évvel ezelőtt még csak baktériumok léteztek a Földön. Ki gondolta volna – hiszen még nem léteztek gondolkodó lények –, hogy az élővilág evolúciója ide fog vezetni. A futurista baktérium olyan élőlényeket vizionált, amelyekben tízbillió sejt egyetlen szuperközösséget alkot, ahol nem kell minden egyes sejtnak minden életfunkciót elvégeznie, szakosítva vannak: „Most úgy áll a helyzet, hogy mindent magunknak kell megcsinálni: tápanyagot találni, megemészteni és kiválasztani a melléktermékeket. Elkerülhetetlen, hogy sejteink összetömrüljenek, és minden egyes sejt specializálja a működését.” Társa hitetlenkedve hallgatja, hogy „a kémiai gradiensek cserélgetése” helyett a kommunikáció új módja fog kialakulni: a hang. „A hang? – kérdezi. – Az meg mi?”

A beszélgetés befejező mondatai:

„– Álljunk csak meg egy pillanatra! Ezek szerint el fogjuk veszíteni alapvető baktérium mivoltunkat?

– Na de valójában semmit sem veszítünk.

– Mindig ezt mondogod, de...

– Nagy előrelépés lesz. Ez a végzetünk. És egyébként is, mindig lesznek ilyen kis baktériumok, mint mi.”

(Harari jóslata a *Homo Deus*-ban: a *Homo sapiens* másodrendű, alávetett lény lesz abban a világban, ahol a technológiával feljavított *Homo Deus* fog uralkodni.)³⁰

Élet a mátrixban

A mesterséges intelligencia – a legújabb technológiákkal együtt (genomika, gén-szerkesztés, nanotechnológia) – paradicsomi állapotokat, digitális mennyországot ígér. Kiküszöböljük a betegségeket, sőt, a halált is. Agyunk letölthető lesz, és virtuálisan halhatatlanok leszünk. Utódainkat optimalizáljuk, tanulásra nem lesz szükség, tudás, tapasztalat közvetlenül betölthető lesz az agyba – sőt, még emlékekkel, kellemes idegen emlékekkel is feltölthető lesz az agyunk. Kedvünkre szabott virtuális valóságban fogunk élni.

De valóban ez kell nekünk? A problémamentes, gondtalan élet? Az evolúció hosszú folyamata során a természetes kiválasztódás nem a boldogságra, hanem a valóságos körülményekhez való igazodás képességére szelektált minket. „Nem tudjuk, hogy miért vagyunk e világon, de azért biztosan nem, hogy boldogok legyünk” – írta Wittgenstein.³¹ Az ember veleszületett, kiolthatatlan szabadságigénye azt is magában foglalja, hogy szabad legyen tévedni, szenvedni, rossz útra tért, saját hibáinkból tanulni. A halál megszüntetése, az öröklét végtelenül unalmas, érdektelenné tenné az életet. Semminek nem lenne tétje, hiszen bármi akárhányszor megismételhető. Ugyanígy, a rutinszerűen, végeláthatatlanul bekövetkező örömök sem jelentenék azt, ami csak ritkán vagy kivételesen történik

³⁰ Yuval Noah Harari: *Homo Deus*.

³¹ Idézi Richard David Precht: *Künstliche Intelligenz und der Sinn des Leben*. 205.

meg.³² Az élet azért ilyen nagyszerű, mert egyszeri és megismételhetetlen. Életünknek a halál ad értelmet.

Mic vogmuc?

Siránkozunk, hogy a digitális jövő, a gépi intelligencia világa emberi mivoltunkat, emberi lényegünket fenyegeti. De ezt anélkül tesszük, hogy világosan látnánk, mik is vagyunk, mi az emberi lényegünk, és milyen jellegű a fenyegetés, amely azt veszélyezteti. Ezt a lényegbevágó kérdést Major Gyöngyi tanulmánya veti fel és elemzi.³³

„Mit jelent élőnek lenni, mi tesz valakit emberivé, mi tekinthető az emberit felülmúlónak? Az emberi faj milyen változásait vagyunk képesek elfogadni? Kinek a szemszögéből tekintünk valamit korrekciónak? Kinek a szempontjából tekintünk valamit »emberinek«? Miféle emberi lényeket akarunk létrehozni? Gépeink és számítógépeink hús-vér másolatát? Vagy mindenségre nyitott tudatú lényeket, olyan szervezeteket, amelyek boldogan fejlődnek valamilyen, még soha nem látott irányba?”

„Az ember identitása éppen nem örök érvényű értékek mentén jön létre, hanem adott kultúra tudati terméke, s ekként meghatározó kérdéssé válnak azok az egymással kölcsönhatásban lévő tényezők, amelyekből integrálódhat az identitás. Az önkonstrukció lehetőségére rátelepedő technológiai determináltságban a MI egyaránt felülírja az esszencialista és a nem esszencialista antropológia felvetését, mégpedig úgy, hogy az ember-kép nem egy újabb értelmezési lehetőséget kap, hanem éppen, hogy kikerül saját értelmezési lehetőségrendszeréből.”

„A MI lényeges újdonsága várhatóan éppen arról szól, hogy kultúrájának konstruálhatósága felett veszíti el hatalmát a létezés aktora. A szingularitás várható következményrendszerében éppen az a koncepció veszíti érvényét, mely szerint »az ember lényege feletti döntés az ember konkrét közreműködése nélkül egy semleges struktúra semleges definíciójában nem található meg, csak történetiségében mint egy mindig újra meghozott döntés«.³⁴ [...] A MI perspektívájában az ember úgy vonódik meg önidentikus rögzíthetőségétől, hogy közben szabadság lehetőségeit is elveszíti. A fő kérdés tehát az, hogy a technikai determinizmus kontextusában milyen lehetőség marad a végső megalapozás (*Letztbegründung*) átmentésére a MI realitásába.”

„Olyan identitás lehetőségét keressük, amely nem a »kultúra és természet« oppozíciójának meghaladhatóságából válik értelmezhetővé, hanem éppen a két entitás oppozícióba állítása helyett azok komplementer viszonyrendszerét, annak értelmezhetőségét keresi – miközben a kultúrát szinte teljességében a technológia

³² A meghalásra képtelen ember gyötrelmeiről lásd Simone de Beauvoir *Minden ember halandó* c. regényét.

³³ Lásd a IV. fejezetben: Az emberi lét-koncepció újraértékelése.

³⁴ Plessner, H., *Macht und menschliche Natur*.

dominánságában értelmezi. Az antropológia vonatkozásában itt már nem a biológia-kultúra ellentétpár a releváns, hanem éppen az ezek szintézisében értelmezett ember az egyik, s az emberről levált gépi/mesterséges kultúra a másik oldalon. Nyitott antropológiát keresünk, mely nem az ember meghaladása mellett érvel, azonban – elsősorban a MI realitása irreverzibilitásának következtében – az emberi mivolt újrafogalmazását mint kikerülhetetlen feladatot látjuk. A technológia és a létezés közötti új egyensúlyt keressük.”

„Olyan identitás kerestetik, amelyben lehetséges az emberi mivolt folyamatként való értelmezése, ami alapján az önmagát »itt és most«-ba abszolutizáló pontszerű énkép önmaga meghaladásának potencialitásával válna, a létezés komplex tér-idő hálózatában fogalmazódhat újra. A kultúra folyamatos megújulásának aktoraként megnyilvánuló létszerep helyett az örök emberi lényeg aktualizációját kifejező ciklikus, de komplexebb struktúrába kibomló emberi megnyilvánulást motiváló létkoncepció meghatározása a feladat. Ezért olyan perspektívát keresünk, amelyben az »örök emberi« respektálása mellett az embert *Homo absconditusnak* (Plessner), azaz »nyitott kérdésnek« is tekinthetjük.”

„Az embert nyitott kérdésként nem az antropológiai eldöntetlenség vonatkozásában értelmezzük, tehát nem Giorgio Agamben *Homo sacer* határlényével rokonítjuk, aki lényegében a »semmi« megtestesítője egyfajta légüres térben, hanem éppen ellenkezőleg, egy olyan személyiségidentitáshoz kötve tárgyaljuk, amely a konvencionális antropológiai struktúráját komplexebb struktúrába ágyazottan képes értelmezni. A prediszpozícióra fókuszálunk, amely az élő/biológiai rendszerként működő létezés mivoltaként képződik le komplexebb struktúrában. Ez a létkoncepció az összekapcsolhatóság nyitott és befejezetlen jellegében képezi le a természet által működtetett rendszerstruktúrát.”

„Azt keressük, hogy a MI realitásában milyen emberképpel biztosítható az, hogy a természeti környezet a digitális jövő 4.0 valóságában is megmaradjon a technológiai dimenzió komplementer tényezőjének; az egzisztenciák kölcsönhatásában keressük a *coexistencia* sajátos ontológiáját.”

Megváltás

E kötet szerzőinek többsége bevallja, tanácstalan a jövőt illetően. Világosan látni véljük, hogy milyen veszélyek fenyegetik az emberi személyiséget, de nem tudjuk, mi a teendő. Az optimista jövőképek nem nyugtatnak meg. Nem látjuk világosan, hogy ki irányítja a történelmet, milyen szerepünk van a jövő formálásában, mitől függ a technológia fejlődése, kézből tartható-e, vagy determinizmussal állunk szemben.

A tanácstalanság abban is megnyilvánul, hogy miközben az egyik szerző Arisztotelész kétértékű logikáját teszi felelőssé azért, amiért idejutottunk, állítván, hogy az arisztotelészi logikából és világtépből logikusan vezet az út a gépig és a mesterséges intelligenciáig, a kötet egy másik szerzője azért sópánkodik, hogy a kiberfilozófusok – felrúgva Arisztotelész duális ontológiáját és logikáját

– megnyitják az utat a nem biológiai entitások szubjektumnak tekintése előtt, és ezáltal ember és gép összeolvadása előtt egyengetik az utat.

De elkeseredésünkben, tanácstalanságunkban idézzük fel az apostolokat. És gondoljunk bele: egy megszállott ember, egy próféta maga köré gyűjtött egy tucat hétköznapi embert (volt köztük halász, ács, vámszedő – a későbbiekben egy sá-torkészítő is csatlakozott), és olyan tanokat kezdett el hirdetni, melyekből kialakult a kereszténység, és ez a vallás kétezer éven át a nyugati civilizáció meghatározó elemévé és tartókövévé vált, sőt, a világ urává tette azt. A nagy fordulat akkor történt meg, amikor ez a jelentéktelen, a szerénységet, az alázatot, a jóságot (azaz mindazt, ami a rómaiak ideáljainak az ellenkezője volt) hirdető új vallás Istene befurakodott a rómaiak diadalmas, hatalmas, gazdag istenei közé, majd egy idő múltán egyeduralmukodóvá vált közöttük. Hogy hogyan történhetett meg mindez, arra különféle magyarázatok vannak. Én a leghihetőbbnek azt tartom, hogy a rómaiak elképedve látták: a keresztények még akkor sem voltak hajlandók megtagadni hitüket, amikor vadállatok elé vetttévén azok szétmarcangolták őket az arénában.³⁵ Ebből arra következtettek, hogy a keresztények istene erősebb, mint az övéik, ha már ilyen hatalmas hittel és elszántsággal ruházza fel híveit.

Mi legyen az új vallásunk? És milyen áldozatokat hozunk? Tagadjuk meg az internet használatát? És – mai ludditákként – törjük össze a laptopokat, az iPadeket meg az okostelefonokat? Mondjunk le az információáramról, melyben oly boldogan tobzódunk? És a kommunikáció szinte végtelenné tágult határaitól, melyet oly önfeledten élvezünk?

És kikkel kell felvenni a harcot? A világ legnagyobb internetes cégeivel és információtechnológiai vállalataival? Melyek már tőkeerőben és gazdasági súlyban minden más szektort és vállalatot leelőznek? És hogyan? Ne internetezzünk, ne fészbukozzunk? Verjük ki gyerekeink kezéből az okostelefonokat? (– Melyeket tőlünk kaptak születésnapjukra.) Ne engedjük őket internetes játékokat játszani?

És hogyan lehet leállítani az e téren folyó nemzetközi versenyt? Amikor ebben a versenyben a nagyhatalmak kormányai a legérdekeltebbek... Az Apple, a Microsoft, az Amazon, az Alibaba, a Tencent működését nemcsak profitéségük motiválja; a mögöttük álló kormányok hatalmi harcának ez a legfontosabb politikai-gazdasági-stratégia-katonai terepe.

És van-e olyan határ, amelyre azt lehet mondani, hogy ezen nem szabad átlépni, mert ha megteesszük, kontrollálhatatlanná válik a helyzet, nem tudunk többé uralkodni a technológia fölött? Nem lett volna szabad szekvenálni a DNS-t? Nem lett volna szabad klónozni Dollyt? Nem lett volna szabad az informatikát ráengedni a mikrobiológiára? A szoftverfejlesztőknek meg kellett volna állni a *deep learning* előtt? Úgy be kellene tiltani minden, a kvantumkomputerek megalkotását célzó kutatást, ahogy korábban az atombomba létrehozását is meg kellett volna tiltani? De Erwin Chargaff osztrák-amerikai biokémikus szerint már a mikroszkóp feltalálása is megengedhetetlen felelőtlenség volt...

³⁵ Hogy kissé oldjam a helyzet tragikumát: egy kisfiú megszólal az arénában: „Nézd anyu, annak a szegény, öreg, fogatlan oroszlánnak egyetlen keresztény sem jutott!” (Márai)

De először is, és mindenekelőtt kellene egy próféta. De honnan akasszunk le a mai világban egy prófétát? És észrevesszük-e majd, ha megjelenik? Nem járunk-e úgy, mint Szakonyi Károly tévévezésbe belefeledkezett hősei az *Adáshibában*?

Én látni vélek egyet: Max Tegmark svéd származású amerikai informatikus személyében. Megítélésem szerint az informatika és a mesterséges intelligencia területén ő a legkompetensebb tudós. Ráadásul 2015-ben mozgalmat indított el „a jó szándékú mesterséges intelligencia” mellett. Hadakozása bármennyire is kilátástalannak tűnik, bízzunk benne és reménykedjünk, hogy sikeres lesz. És kövessük, csatlakozzunk hozzá! Kedvező jel, hogy már az internetóriások is érzékelik a veszélyt, és megpróbálnak tenni ellene. 2018-ban a Google közreadta mesterséges intelligenciával kapcsolatos irányelveit, többek között azt, hogy a MI-rendszerek társadalmilag hasznosak és biztonságosak kell hogy legyenek. Hasonló követelményeket fogalmazott meg a Microsoft és a Facebook.³⁶

A kereszténység esetén kívül is megtörtént már, hogy egy szellemi törekvés megváltoztatta a világot. (Nevezetesen éppen azt a világot, amelyet a kereszténység hozott létre.) Az 1700-as évek francia enciklopédistái képesek voltak olyan mozgalmat elindítani, a felvilágosodást, mely megrendítette a keresztény egyház hatalmát, és mára már Európában szinte eljelentéktelenítette azt. Tehát nem példa nélküli, hogy emberek egy csoportja képes volt tudatosan megváltoztatni a történelem menetét. (Sőt, egy-egy ember! A Buddha, Konfuciusz, Jézus vagy Mohamed. Rossz példának Marxot is említhetnénk.) Amikor a változtatás nem hatalmi harc következményeként, „melléktermékeként”, nem országok közötti háborúk vagy osztályok közötti harcok eredményeként történt meg, és nem is a technológiai fejlődés eredményezte; hanem a világ, az ember jobbításának szándéka.

Közbeavatkozás nélkül a világ menete jó előre láthatóan azon a pályán fog haladni, amelyet Kurzweil megjósolt és az univerzum evolúciójának nevezett. Ebben a történetben a *Homo sapiensre* való hivatkozás/emlékezés úgy fog majd kezdődni, hogy: „Volt egyszer egy ember...” És mi ezen az univerzális, elkerülhetetlennek tűnő folyamaton szeretnénk változtatni...

A hivatkozott források

- De Beauvoir, Simon: *Minden ember balandó*. Ford. Görög Lívia. Miskolc, Bíbor Kiadó, 2003.
 Harari, Yuval Noah: *Sapiens*. Budapest, Animus, 2015.
 Harari, Yuval Noah: *Homo Deus*. Budapest, Animus, 2017.
 Heisenberg, Werner: *A rész és az egész*. Budapest, Gondolat, 1978.
 Kiss Károly: *Ki akar itt Superman lenni?* Budapest, L'Harmattan, 2019.
 Kurzweil, Ray: *A szingularitás küszöbén*. Budapest, Ad Astra és web, 2013.
 Plessner, H.: Macht und menschliche Natur. Ein Versuch zur Anthropologie der geschichtlichen Weltansicht. In Dux, Günter – Marquard, Odo – Ströker, Elisabeth – Schmidt, Richard W.

³⁶ *The Economist*, 2020. június 13.

- Wetterer, Angelika – Zemlin, Michael-Joachim (Hrsg.): *Gesammelte Schriften*. 10 Bde. Frankfurt am Main, Suhrkamp, 2003. 135–234.
- Precht, Richard David: *Künstliche Intelligenz und der Sinn des Lebens*. München, Goldmann, 2020.
- Rees, Martin: *Üvött az utolsó óránk?* Budapest, Athenaeum, 2004.
- Tegmark, Max: *Élet 3.0 – Embernek lenni a mesterséges intelligencia korában*. Budapest, HVG Könyvek, 2018.
- The Economist* June 13th 2020. „Steeper than expected. Artificial intelligence and its limits. Technology Quarterly”.
- Lev Tolsztoj: *Háború és béke I–II*. Ford. Makai Imre. Budapest, Új Magyar Könyvkiadó, 1954