

Emberi értelem, mesterséges intelligencia - a társadalom értelmi felépítettségének változásai

A fizikai és a biológia létszféra fölé emelkedő társadalmi lét a milliárd számban élő embertömegeinek biológiai létén alapul, de a sajátlagosan társadalmi ebben az *értelmi* létezés egyre markánsabb kibomlása. Ez a fizikai-biológiai létfeltételeken alapulás az elmúlt évezredek társadalmainak fejlődése során - a kezdeti teljes beágyazódás után - egyre inkább "elvékonyodott", és különösen az elmúlt egy-két évszázad során háttérbe szorult egy sor aspektusban. Így az írás egyre szélesebb használata az emberi érintkezésben - az írástudás tömegessé válásával - az emberi kapcsolatok közvetlen fizikai jelenlétre ráutaltságát csökkentette, az újságok rendszeres olvasásának kialakulása a XX. századtól ezt csak fokozta, majd a telefon megjelenése még tovább csökkentette a fizikai jelenlétre ráutaltságot az emberi kapcsolatokban és a társadalom szerveződésében, melyet a televíziózás mindennaposá válása követett, majd legutóbb az internetre alapozódó társadalmi érintkezések tömegessé válása jött létre, mely a legnagyobb mértékben háttérbe tolta a fizikai-biológiai beágyazottságra épülést a társadalmak szerveződésében. Ily módon egyre inkább csak *végső fokon ráépülést* jelent mára a társadalmak mind sűrűbb értelmi szerveződésében a fizikai-biológiai léte épülése, és az értelem az emberi-társadalmi evolúcióval mintegy kiszabadul az alsóbb létszférák fogságából.

Ez az elméleti belátás az 1880-as évek végén jelent meg, és felhasználva a platóni dichotómiát az érzékileg létező és az ideálisan létező között Wilhelm Dilthey volt az, aki elkezdte az 1800-as évek vége felé a természeti létezők és a szellemi létezők elválasztásával a fizikai-biológiai világ fölé emelkedő társadalmi világ sajátos anyagának, az *értelmi létnek* a feltárását. Vele párhuzamosan, illetve némileg később Edmund Husserl is a társadalmi világ értelmi felépítettségének sajátosságait dolgozta ki, ahogy a főbb vonalat illetően Max Weber is a társadalmiság értelmi felépítettségét vallotta. Alfred Schütz 1932-ben Husserl és Weber elméleti kiindulópontjainak integrálásával megírta "Der sinnhafte Aufbau der sozialen Welt" (*A társadalmi világ értelmi felépülése*) c. könyvét, melyből később a társadalom értelmi puhaságát és formálhatóságát középpontba emelő tudásszociológia főiránya alakult ki. Ugyanekkor, de részben más irányból Nicolai Hartmann is a társadalmiság értelmi anyagát emelte a középpontba a *das geistige Sein* és a természeti világ valóságsszféráinak elkülönítésével. Mindezeket aztán 1960-as évektől Niklas Luhmann foglalta össze egy egységes társadalomelméletben. („Az értelem, mint a szociológia alapkategóriája” c. írását már 1986-ban beválogattam a Szociológiai Füzetek 42. számaként megjelent első hazai Luhmann-kiadásba).

Így felfogva a társadalmiság felépülésében az értelem szerepét, a gépi értelem és a mesterséges intelligencia egyre fokozódó megjelenési formái az elmúlt években a korábbi társadalomelméleti anyagaim gondolati fonalának újrafelvételére ösztönöznek. Egy sor leírás a gépi értelem és a mesterséges intelligencia emberi és társadalmi ellenőrzés alóli kiszabadulását veti fel, és viharos fejlődésük az utóbbi években hihetővé teszi fejlettségüknek az emberi értelem sokszorosára növekedését. Mivel az elmúlt néhány évben két összegzés is megjelent magyar nyelvre lefordítva az itt

¹ Egyetemi tanár (ELTE ÁJK), alkotmánybíró.

várható fejleményekről, így ezek gondolati összefoglalásával kezdem az elemzést. Az egyik Ray Kurzweil "A szingularitás küszöbén" c. könyve, a másik Nick Bostrom "Szuperintelligencia" c. könyve. E két összefoglalást azért is érdemes párhuzamosan megismerni, mert a gépi értelem és a mesterséges intelligencia jövőbeli veszélyeinek egymással szembenálló felbecslését jelentik. Abban egyetértenek, hogy a következő évtizedekben ezek mindenképpen az emberi értelem fölé nőnek, és ennek ellenőrzésétől elszakad a mesterséges intelligencia. Itt az eltérés csak abban van közöttük, hogy ezt Kurzweil már az 2040-es évek elejére teszi, míg Bostrom a 2080-as évekre prognosztizálja ezt. A nagy szembenállás inkább abban van, hogy eltérő veszélyességet látnak az ennek következtében létrejövő helyzetben. Kurzweil ezt letudja azzal, hogy tényleg vannak ugyan veszélyek, de ezeket kezelni lehet majd, és a veszélyhelyzet magától megoldódik ugyanúgy, ahogy eddig is mindig megoldódtak a gépi fejlődéssel létrejött technológiai veszélyek. Ezzel szemben Bostrom a legnagyobb és szinte az ember létezés eltűnését okozó veszélynek látja az emberi értelem fölé növekvő, és annak ellenőrzésétől elszakadó mesterséges intelligencia által létrehozott helyzetet.

A következőkben először olvasónaplószerűen kivonatolom Kurzweil és Bostrom könyveinek főbb állításait, majd az ezt követő részben megpróbálom értékelni ezeket, és állást foglalni a az emberi értelem fölé növekvő gépi értelem és mesterséges intelligencia főbb kihatásait illetően. A tanulmányt az agy és az elme működéséről szóló újabb Kurzweil-könyv kivonatolásával zárom - amely 2012-ban jelent meg angol nyelven -, de mivel ez csak megalapozó jellegű a szűkebb mesterséges intelligencia témáját illetően, így függelékben helyezem el ezt a tanulmány végén.

I. Olvasónapló

I.1. Ray Kurzweil "A szingularitás küszöbén"

A szingularitás küszöbén c. könyv 2005-ben jelent meg, és magyar fordítása 2013-as. Ebben Kurzweil három technológiai - a biotechnológia, a nanotechnológiai és a robottechnológiai - forradalom összefonódó hatásaiként az emberi történelem eddigi menetének megszűnését, és egy új korszak eljövételét prognosztizálja, és ennek nevéként használja a "szingularitás" elnevezést. Ez az elnevezés több szerzőnél is megjelent az utóbbi években, és röviden úgy lehet leírni e korszak sajátosságát – és kezdetét -, hogy ez akkor áll be, ha a mesterséges értelem eléri és meghaladja az emberi értelem fejlettségét, és ezzel több-kevesebb fokban el is szakad az emberi ellenőrzés alól. Kurzweilnél ez annyiban egészül ki még egy további aspektussal, hogy ő azt is kiemeli, hogy amíg emberi közreműködéssel történik még a mesterséges intelligenciával egyre nagyobb mértékben átítatott társadalmi napi élet menete, addig a milliószeres gyorsaságra is képes mesterséges intelligencia fejlődése és a mind újabb megoldások létrehozása az emberi értelem lassúságára szabottan visszafogott marad. Am ha végbemegy a mesterséges intelligencia által vezérelt folyamatokból az emberi közreműködés teljes kikapcsolása, akkor a kényeszerű lassítás eltűnésével rögtön ezerszeres-milliószeres gyorsaságra ugrik fel a mesterséges értelem által tervezett, kigondolt és létrehozott folyamatok változása. Néhány óra, napok alatt annyi változás történik ezen időpont után, mint korábban sok év alatt, és ennek megértéséből, illetve ellenőrzési lehetőségétől az emberi értelem már az első órák után végleg lemarad. Ez a szingularitás eljövetele, és ezt Kurzweil a 2040-es évek elejére teszi.

(*A technológiai fejlődés exponenciális jellege*) Az ember spontán módon hajlik arra, hogy lineáris vonalban gondolja végig a jövődő változásait, és ahogy az eddigi tapasztalatai mutatják, meghosszabbítva becsüli meg azt, hogy milyen hosszú idő kell egy, a múltban végbement folyamat duplájára növekedéséhez a jövőben. A technológiai fejlődés azonban Kurzweil szerint a hosszú tapasztalatok szerint exponenciális és nem lineáris gyorsassággal megy előre: “A modelljeim azt mutatják, hogy minden évtizedben megkétszerezzük a paradigmaváltás sebességét. (...) Így gyorsult fel a XX. század a mai fejlődési sebességre. A 2000. évi sebességgel számolva - állandó ütemben - a múlt század egésze alatt tapasztalt fejlődéshez csupán húsz évre lett volna szükség. Ezt a húszéves fejlődést aztán tizennégy év alatt 2014-re ismételjük meg, majd mindössze hét év alatt ismét. Másképp kifejezve, nem százéves technológiai fejlődést fogunk tapasztalni a XXI. században, hanem húszezer évnit, azaz ezerszer nagyobb, mint a teljes XX. századé volt.” (Kurzweil 2013: 35). Megítélése szerint ebből a félreértésből adódnak sokszor pesszimista prognózisok pl. az agy kutatás jövőbeli fejlődésének sebességére, mert látják, hogy például egy fonálféreg néhány száz neuronból álló kis agyának számítógépes szimulálására sok év kellett, így az emberi agy sokmilliárdos neuronhálózataik illetve közöttük a billiószoros szinapszisok (neuron-kapcsolódások) szimulálására ezer évet is kevésnek gondolnak. Ám ami tíz év alatt lineárisan számolva csak tíz év fejlődést jelent, az exponenciálisan számolva - évenként állandóan duplázódó ütemet feltéve – több mint ezer év fejlődést hoz létre. “Például nemrég vettem részt egy molekuláris gyártás kivitelezhetőségéről tartott vitában, és egy Nobel-díjas résztvevő azzal hárította el a nanotechnológiával kapcsolatos biztonsági aggályokat, hogy “még száz évig nem fogunk önreplikáns, nanorészecskékből felépülő entitásokat (molekuláris darabról-darabra létrehozott eszközöket) látni.” Száz év megalapozott becslés, sőt egybeesik az én becsléseimmel is, arról hogy mekkora technológiai fejlődés kell ennek a bizonyos mérföldkönek az eléréséhez a *mai fejlődési sebességgel* (ami ötször nagyobb, mint a XX. század átlagos fejlődési sebessége). Ám mivel minden évtizedben megduplazzuk a fejlődési sebességünket, a mai fejlődési sebességgel számított egy évszázadnyi fejlődés mindössze huszonöt naptári év alatt fog bekövetkezni.” (36.p.)

(*A három, egymást átfedő forradalom*) “A XXI. század első felét három, egymást átfedő forradalom fogja jellemezni - a genetika, a nanotechnológia és a robotika forradalma. Ezek fogják elhozni a korábban említett ötödik korszakot, a szingularitás kezdetét. Ma a “G” forradalom (genetika) korai szakaszában vagyunk. Az élet mögött rejlő információs folyamatok megértése révén kezdjük megtanulni biológiánk átprogramozását, hogy gyakorlatilag megszüntessük a betegségeket, drámaian meghosszabbítsuk az életet. (...) Az emberek azonban “másodrendű robotok” maradnak, ami azt jelenti, hogy a biológia soha nem lesz képes felvenni a versenyt azzal, amit a technológia révén tudunk majd létrehozni, amint teljes mértékben megértjük a biológia működési elveit.

Az “N” forradalom (nanotechnológia) lehetővé fogja majd tenni, hogy áttervezzük, és molekuláról molekulára átépítsük a testünket, az agyunkat és a világot, amelyben létezőnk, messze meghaladva a biológiai korlátokat. A legnagyobb közelgő forradalom az “R” (robotika); az emberi szintű robotok intelligenciája a miénkből fog származni, de messze meghaladja az emberi képességeket. Az “R” képviseli a legjelentősebb átalakulást, mivel az intelligencia a leghatalmasabb “erő” a világegyetemben. (...)

Noha mindegyik forradalom meg fogja oldani a korábbi átalakulások nyomán keletkezett problémákat, új veszélyeket is magával hoz. A “G” leszámol a betegségekkel és az öregedéssel, de egyben alapul szolgál új, biológiailag tervezett vírusfenyegetéseknek is. A teljes mértékben kifejlett “N” képes lesz megvédeni minket a biológiai veszélytől, de megteremti a saját önszaporodó fenyegetését, ami sokkal nagyobb lesz, mint bármilyen biológiai veszély. A teljesen kifejlett “R” révén meg fogjuk tudni védeni magunkat ezektől a fenyegetésektől, de mi fog megvédeni minket a miénket meghaladó, beteges intelligenciától? Van stratégiám ezeknek a problémáknak a megoldására, a 8. fejezet végén fejtem ki részletesen. Ebben a fejezetben azt vizsgálom meg, hogyan fog kibontakozni a szingularitás a három, egymást átfedő forradalom, a “G”, az “N” és az “R” révén.” (294-95.p.)

(A *humán géntechnológia lehetőségei* 296-323.p.) “Az élet összes csodája és betegségek összes nyomorúsága mögött meglepően tömör információs folyamatok, lényegében szoftverek rejlenek. Az egész emberi genom egy szekvenciális bináris kód, amely mindössze nyolcszázmillió byte információt tartalmaz. Mint korábban említettem, amikor hagyományos tömörítési eljárásokkal megszüntetjük óriási redundanciáját, mindössze harminc-százmillió byte marad, ami megfelel egy mai átlagos számítógépes szoftver méretének. (...) Ez a gépezet lényegében egy önmagát replikáló, nano nagyságrendű replikátor, amely felépíti mindazokat a bonyolult hierarchiájú struktúrákat, és egyre összetettebb rendszereket, amelyekből egy élőlény áll.

(...)

Már ma is elegendő információ áll rendelkezésünkre a betegségek és az öregedési folyamatok olyan mértékű lelassításához, hogy nemzedékek hozzám hasonló képviselői jó egészségben megérjék a biotechnológiai forradalmat. (...) Aubrey de Grey, a Cambridge Egyetem genetikai tanszékének tudósa energikus és nagy tudású hirdetője annak, hogy a biológia mögött rejlő információs folyamatok megváltoztatásával meg lehet állítani az öregedést. (...) De Grey “tervezett, elhanyagolható fokú öregedésnek” nevezi a célját – meg akarja akadályozni, hogy a test és az agy az idő múlásával egyre törékenyebb és betegebb legyen.

(...)

A közelmúltban kifejlesztett géntechnológiai eljárásoknak köszönhetően annak a határán állunk, hogy befolyásolni tudjuk a génexpressziót. A génexpresszió az a folyamat, amely révén adott sejtalkotórészek (konkrétan az RNS és a riboszómák) az adott genetikai mintát követve fehérjéket hoznak létre. Noha az emberi test minden sejtje tartalmazza a test génjeinek összességét, az adott sejt, például a hámsejt vagy egy hasnyálmirigy-szigetsejt, csak a genetikai információ rá vonatkozó töredék részéből kapja meg a jellemzőit. (...) A génexpressziót peptidek (maximum száz aminosavból áll molekulák) és rövid RNS-szálak szabályozzák. Sok, új jelenleg fejlesztés alatt álló terápia azon alapul, hogy a peptidek manipulálásával vagy kikapcsolatja velük a betegséget okozó gének expresszióját, vagy bekapcsolatja a kívánatos géneket, amelyek egyébként nem expresszálódnának egy adott sejt típusban.

(...)

Egy másik fontos támadási irány a sejtjeink, szöveteink, sőt akár teljes szerveink újránövesztése és testünkbe ültetése sebészeti beavatkozás nélkül. Ennek a “terápiás klónozás” technikának az egyik fő előnye az, hogy ezeket az új szöveteket és szerveket a saját sejtjeinkből leszünk képesek létrehozni, amiket a kialakulófélben levő fiatalító orvoslás révén már fiatalabbakká tettünk. Például képesek leszünk új szívsejteket létrehozni hámsejtekből, és a vérkeringésen keresztül bejuttatni őket a

szervezetbe. Idővel a meglévő szívsejtek helyét átveszik az új sejtek, és kapunk egy megfiatalított, "új" szívet, amit a saját DNS-ünk felhasználásával hoztunk létre.

(...)

A szomatikus génterápia a nem reproduktív sejtek génterápiája. Ez a biomérnökség Szent Grálja, ami lehetővé fogja tenni, hogy megváltoztassuk a sejtben lévő géneket azáltal, hogy új DNS-sel "fertőzzük meg" a sejtet, s ezzel gyakorlatilag új géneket hozunk létre. Az emberek génkészletének ellenőrzését gyakran összekapcsolják azzal az elképzeléssel, hogy befolyásoljuk az új nemzedékek, a "tervezett csecsemő" tulajdonságait. A génterápia igazi ígérete azonban génjeink születés utáni, felnőttkorban történő megváltoztatása, Blokkolhatjuk a nemkívánatos, betegségekre hajlamosító géneket, vagy újakat hozhatunk létre, amelyek lelassítják, vagy akár megfordítják az öregedés folyamatát.

(...)

A rák és az öregedés szempontjából egyaránt kulcsfontosságúak a telomerszakaszok, a kromoszómák végén található, ismétlődő DNS-szekvenciák. Akárhányszor osztódik egy sejt, ez a szakasz eggyel rövidül. Amikor egy sejt annyiszor osztódott, hogy a telomere egészen elfogy, nem lesz képes tovább osztódni, és elpusztul. Ha meg tudnánk állítani ezt a folyamatot, a sejtjeink örökké élnének. Szerencsére újabb kutatások megállapították, hogy ennek elérésére, egy egyszerű enzimre (telomerázra) van szükség. A trükk abban rejlik, hogy úgy adagoljuk a telomerázt, hogy ne okozzon rákot. A rákos sejteknek van egy génjük, mely telomerázt termel, és ez gyakorlatilag halhatatlanná teszi őket, a végtelenségig osztódhatnak. A rák elleni harchoz tehát az is hozzátartozik, hogy meggátoljuk a rákos sejteket a telomeráz termelésében. (...) Ha klónozott, meghosszabbított telomerú, kijavított DNS-ű sejteket juttatunk a szervezetbe, akkor beépülnek a régi sejtek közé. Ha ezt a kezelést újra és újra megismételjük, a szóban forgó szervet végül túlnyomórészt fiatal sejtek fogják alkotni. Egyébként is rendszeresen lecseréljük sejtjeinket, akkor miért ne cseréljük őket megfiatalított sejtekre a rövid telomerú hibásak helyett? Semmi sem indokolja, hogy miért ne ismételhettünk ezt a folyamatot a testünk összes szervével, miért ne változnánk egyre fiatalabbakká. (...) Amerikai és norvég tudósok a közelmúltban sikeresen átprogramoztak májsejteket hasnyálmirigysejteké. Egy másik kísérletsorozatban emberi hámsejteket alakítottak át úgy, hogy azok az immunrendszer sejtjeinek és az idegsejteknek számos tulajdonságával rendelkezzenek.

(Az éhezés felszámolása a világban) A klóntechnológia még az éhezésre is talál megoldást: húst és egyéb fehérjeforrásokat gyárakban is elő lehet állítani, állatok nélkül az állati izomszövet klónozásával. Az előnyei közé tartozna a rendkívül alacsony költség, a mai, nagyüzemi körülmények között termelt húsokban lévő féregirtó szerek és hormonok hiánya, a nagymértékben csökkenő környezeti hatás (a nagyüzemi állattartáshoz képest), a jobb tápanyagprofil, és az állatok szenvedésének megszűnése. Akárcsak a terápiás klónozás esetében, nem teljes állatokat klónoznánk, hanem csak testrészeket vagy húst, Lényegében az összes hús - több million tonna - ugyanattól az állattól származna.

(Nanotechnológiai forradalom 325-368 p.) "Bármilyen fontos is lesz a fent ismertetett biotechnológiai forradalom, a módszerek kiérlelése után szembesülnie kell a biológia saját korlátjaival. (...) A nanotechnológiai forradalom viszont végleg képessé tesz minket arra, hogy molekuláról molekulára tervezzük át, építsük újra testünket, az agyunkat és a minket körbevevő világot egyaránt. Ez a két nagy forradalom nagyrészt

átfedi egymást, de a nanotechnológia teljes megvalósulása nagyjából egy évtizeddel van lemaradva a biotechnológiai forradalom mögött.

(...)

A nanotechnológia konceptuális alapját Neumann János fektette le még az ötvenes évek elején, amikor megalkotta az egyetemes számítógéppel kombinált egyetemes konstruktoron alapuló, önsokszorosító rendszer egyik modelljét. (...) Eric Drexlerre maradt, hogy megteremtse a nanotechnológia modern tudományágát a nyolcvanas évek közepén írt disszertációjában. Drexler leírta Neumann kinematikus konstruktorát, és alkatrészeiként atomokat és molekulákat adott meg. Az 1992-es *Nanosystems - Nanorendszerek* - című könyvében lefektette a nanotechnológia alapjait, amelyekre a mai napig építkezünk. Drexler molekuláris összeállítója (assembler) szinte bármit képes létrehozni. Számítógép biztosítja az összeállítási folyamat irányításához szükséges intelligenciát; az utasítás-architektúra: egyetlen adattároló egyszerre rögzíti és továbbítja az utasításokat több billió molekuláris méretű összeállítónak (melyek mindegyike rendelkezik saját, egyszerű számítógéppel) Az adatokat egy központból sugárzó architektúra egy kulcsfontosságú biztonsági problémát is megold: az önsokszorosítási folyamatot le lehet állítani, ha irányíthatatlanná válna, elég csak a központi forrását kikapcsolni. Az önsokszorosításban rejlő veszélyeket figyelembe véve a Foresight Intézet (Drexler Intézete) által javasolt etikai standardokban szerepel a korlátolatlan önsokszorosítás tiltása, különösen természetes környezetben.

(...)

A molekuláris gyártórendszereket irányító szoftver tervezése maga is szinte teljességgel automatizálva lenne, akár csak napjainkban a mikrochipek terve. A chiptervezők nem határozzák meg a milliárdnyi vezető és egyéb alkatrész mindegyikének pontos helyét, inkább csak a konkrét funkciókat és tulajdonságokat határozzák meg. A számítógépes tervezőrendszerek aztán maguk fordítják át ezt az információt a chip végső tervezetébe. (...) Az összeállító úgy hozná létre a molekuláris robotokat, hogy először legyárt belőlük egy kisebb mennyiséget, aztán ezekkel a robotokkal hozat létre további robotokat, míg csak el nem éri a kívánt mennyiséget. (...) A folyamat a biológiai rendszerek génkifejezésével (expressziójával) lenne analóg. Bár minden egyes sejtben ott van minden egyes gén, csak az adott sejtípus számára pontosan szükségesek kerülnének kifejezésre.

(...)

A molekuláris összeállító megvalósíthatóságára maga az élet a végső bizonyíték. Minél alaposabban átlátjuk az életfolyamatok információs alapját, annál több olyan ötletet fedezünk fel, amik az általános molekuláris összeállító tervezési követelményeire is alkalmazhatók. Nézzük csak meg, a biológia miként oldja meg a Drexler-összeállító tervének a kihívásait! A riboszóma egyszerre képviseli a számítógépet és az építő robotot. Az élet nem központosított adattárolót használ, hanem minden egyes sejtet ellát a teljes kóddal. Az élet adattárolói természetesen a DNS-szálak, a kromoszómákon elhelyezkedő egyes génekkel. Az utasítások kitarakását (az adott sejtípus számára nem szükséges gének blokkolását) a rövid RNS-molekulák és a génkifejezést vezérlő peptidek irányítják. (...) A biológia éppúgy gondosan ellenőrzött forrás és üzemanyagkészlettel dolgozik, melyeket az emésztőrendszerből nyer. (...) A nanoalapú sokszorosítók nagyságrendekkel gyorsabbak és erősebbek bármely biológiai rendszerénél (...) A nanocsövek igen sokoldalú szerkezeti összetevőnek bizonyultak. A Lawrence Berkeley Egyetem Nemzeti Laboratórium tudósai nemrégiben nanocsövekből készült futószalagot mutattak be. A nanoméretű futószalag segítségével parányi indiumrészecskéket szállítottak egyik helyről a másikra.

(...)

A General Dynamics által a NASA számára elvégzett kutatás igazolta a nanoméretű, önsokszorosító gépek megvalósíthatóságát. A kutatók számítógépes szimulációk segítségével bizonyították, hogy a kinematikus sejtautomata névre hallgató, újrakonfigurálható molekuláris modulokból épített, molekuláris pontossággal felépülő robotok képesek önmaguk sokszorozására. (...) A müncheni Ludwig Maximilian Egyetem egy kutatócsoportja olyan "DNS-kezet" épített, amely képes kiválasztani egy adott fehérjét a többi közül, kötést létrehozni vele, majd parancsra eloldani. A molekuláris tárgyak irányított megragadása és elengedése újabb olyan tulajdonság, ami igen fontos a molekuláris nanotechnológiai összeállítás során. (...) A DNS-ből épített nanoméretű szerkezetek különösen lenyűgöző példája az a parányi kétlábú robot, amely mindössze tíz nanométer hosszú lábakon jár. A nanorobot a New York-i Egyetem két kutatója, Nadrian Seeman és William Sherman projektje, és úgy mozog, hogy eloldja a lábát a pályáról, előrelép, majd visszaköti lábát a pályára. Ez a projekt hatásosan demonstrálja, hogy a nanoméretű gépezetek képesek manőverek elvégzésére. (...) A nanorobotok tervezésének alternatív módszere az, ha a természettől tanulunk. A baktériumok természetes nanorobot méretű objektumok, és képesek a mozgásra, az úszásra és a folyadékok pumpálására. Linda Turner, a harvardi Rowland Intézet kutatója a baktériumok különféle feladatok elvégzésére alkalmas vékony karjaira (az ún. fimbriákra) fókuszál, és a baktériumok a karok segítségével szállítanak más nanoméretű tárgyakat, és keverik össze a folyadékokat. Egy másik megközelítésben a baktériumoknak csak a részeit használnák fel. Minthogy a baktériumok eleve számos eltérő funkciót ellátni képes nanoméretű rendszerek, a kutatás végső célja nem más, mint a baktériumok működésének visszafejtése, hogy az onnan ellesett tervezési elveket felhasználhassuk a saját nanoterveinkben is.

(...)

Különösen izgalmas az az alkalmazás, ami a nanorészecskék irányításával a test adott pontján képes kezelést végezni. A nanorészecskék a sejtfalakba kormányozhatják a gyógyszereket, át a vér-agy-gáton. A montreali McGill Egyetem kutatói egy 25-45 nanométeres szerkezettel bíró nanopirulát mutattak be. A nanopirula elég kicsi ahhoz, hogy átjusson a sejtfalon, és egyenesen a sejten belül megcélzott struktúrához juttassa el a gyógyszert. A japán tudósok száztíz aminosavból álló nanoketreceket hoztak létre. A ketrecekben gyógyszer-molekulák találhatóak. Minden egyes nanoketrec felszínéhez egy-egy olyan peptid kapcsolódik, amely az emberi test megadott pontjaival létesít kötést. Az egyik kísérletben a tudósok olyan peptidet használtak fel, amelyik az emberi májsejtek egy meghatározott receptorával alakít ki kötést. (...) Nanoméretű csomagokat lehetne arra tervezni, hogy a bennük rejlő gyógyszert megvédjék az emésztőrendszerrel, és adott helyre szállítsák, aztán összetett módon adagolják, azt sem kizárva, hogy az esetleg a testen kívülről érkező utasításoknak engedelmessé válnak. A floridai Alachuában található Nanotherapeutics által kifejlesztett, alig pár tucat nanométer vastag, biológiailag lebontható polimer ezt a megközelítést használja. (...) A Texasi Egyetem kutatói olyan nanobot méretű cellát hoztak létre, amely egyenesen az emberi vérben található glükóz-oxigén reakció révén állít elő elektromos áramot. A "vámripírbot" néven emlegetett cella kellő energiát termel a hagyományos elektronikus szerkezetek működtetéséhez, és talán fel lehet használni a jövőben a vérben keringő nanobotok működtetéséhez is.

(...)

"A nanogyógyászati beavatkozások nettó hatása a biológiai öregedés minden aspektusának leállítására mellett kiterjed arra is, hogy a páciens testének adott biológiai állapotát – a "korát" – az általa kívánt új biológiai állapotra - "új életkorra" -

cseréljük. Ezáltal elvágja a naptár szerint eltelt idő és a biológia egészség közti kapcsolatot. Jó néhány évtized múlva mindennaposak lesznek az efféle beavatkozások. Az éves vizsgálatok és tisztítókúrák, és néhány nagyobb javítás elég lesz ahhoz, hogy a biológiai korunkat évente egyszer visszaállítsák arra a nagyjából azonos fiziológiai korra, amit választunk, Persze balesetben még mindig meghalhatunk, de legalább tízszer olyan hosszú ideig élünk, mint jelenleg.” (Robert A. Freitas Jr.)

Robert Freitas, a nanotechnológia elméletének egyik úttörője és a nanogyógyászat (azaz a biológiai rendszereink molekuláris méretű konstrukciók segítségével történő átkonfigurálásának) egyik fő támogatója olyan robotokat tervezett, melyek lecserélhetnék az emberi vértesteket, és több százszor vagy több ezerszer hatékonyabbak lehetnének biológiai megfelelőiknél. Freitas respirocitái (robotvörsvértestei) segítségével egy futó lélegzétvétel nélkül tizenöt percen át bírná az olimpiai sprintet. DNS-javító robotjai képesek lennének kijavítani a DNS-másolási hibákat, vagy esetleg szükség szerint megváltoztatni a DNS-t. (...) Bár Freitas koncepcióit még pár évtizedig nem tudjuk megvalósítani, máris jelentős előrelépések történtek a véráramba juttatható készülékek terén. A chicagoi Illinois-i Egyetem kutatói például patkányokban sikeresen gyógyították az 1-es típusú cukorbetegséget hasnyálmirigy-szigetsejteket tartalmazó, nanotervezett gépezettel. A készüléken hét nanométer átmérőjű pórusok engedik ki az inzulint, de nem engedik be azokat az antitesteket, amelyek elpusztítanák az inzulint termelő sejteket.

(Robotika forradalma - erős mesterséges intelligencia 374-426 p.) “Míg a GNR “R”-je a robotikát jelenti, a valódi kérdés az erős mesterséges intelligencia (az emberi intelligenciát meghaladó intelligencia) létrejötte. A robotika hangsúlyozása csak azért érdekes, mert az intelligenciámnak testre van szüksége, egy fizikai jelenlétre ahhoz, hogy kapcsolatba léphessen a világgal. Nem értek egyet azonban a fizikai jelenlétnek a hangsúlyozásával, mert úgy hiszem, a központi probléma az intelligencia. (...) Az emberi tudás webre vándorlásával a gépek képesek lesznek olvasni, megérteni és összegezni minden emberi-gépi információt. Az emberi agy merev architektúráját tekintve nincsen kapacitásunk (vagy időnk) minden, egyre specializálódó területen a legmagasabb szintű képességet kifejleszteni és alkalmazni. A nem biológiai intelligencia ezzel szemben képes lesz elérni és meghaladni az emberi képességek csúcspontját az élet minden területén.

Az erős mesterséges intelligencia teljes nanotechnológiához vezet majd (molekuláris összeszerelők formájában, amelyek az információt fizikai terméké alakítják). Azonban azt hiszem, hogy a teljes MNT (molekuláris nanotechnológia) az erős mesterséges intelligencia előtt fog megszületni, habár csak néhány évvel (a nanotechnológia 2025 körül, az erős mesterséges intelligencia 2029 körül). Amint a mesterséges erős MI elkezd sok másik mesterséges intelligenciát gyártani, az utóbbiak már képesek lesznek megérteni és továbbfejleszteni a saját felépítésüket, és így gyorsan egy sokkal nagyobb kapacitású, sokkal intelligensebb mesterséges intelligenciává fejlődni, és ez a körforgás akármeddig folytatódhat. Az egyes ciklusok nemcsak, hogy egy sokkal intelligensebb mesterséges intelligenciát hoznak létre, de mind kevesebb időbe kerülnek majd, mint az azt megelőző ciklusok. Amint megalkotjuk az első gépet, amely átmegy a Turing-teszten (2029 körül), az azt követő időszak a konszolidációról fog szólni, amikor is a nem biológiai intelligencia sebes fejlődésbe kezd. Azonban a szingularitással járó soha nem látott expanzió, melyben az emberi intelligencia megmilliárdszorozódik, nem történik meg a 2040-es évek közepéig.

(A mesterséges intelligencia eszköztára: neurális hálók és genetikai algoritmusok) A kulcs az a neurális hálóhoz, hogy meg kell tanulnia a céltárgyát. Mint az emlősgyagok, amelyekről nagyjából modellezték, a neurális hálók kezdetben ostobák. A neurális háló tanítója - aki lehet egy ember, egy számítógépprogram vagy talán egy másik, érettebb neurális háló, amely már megtanulta a leckéjét – jutalmazza a tanuló neurális hálót, amikor jó eredmény ad, és megbünteti, amikor rosszat. Ezt a visszajelzést a tanuló neurális háló arra használja, hogy beállítsa az interneurális kapcsolatok erősségét. A jó válasszal összefüggő kapcsolatok erősödnek, a rossz választ adók gyengülnek. Idővel a neurális háló átszervezi magát ahhoz, hogy tanítás nélkül is a jó választ adjon. A kutatások bebizonyították, hogy a neurális hálók még a megbízhatatlan tanárokkal is képesek megtanulni a leckét. Ha a tanár csak hatvan százalékban tudja a helyes választ, a tanuló neurális háló még abból is tanul. Egy erős, jól megtanított neurális háló széles körű emberi mintafelismerést képes emulálni. A többrétegű neurális hálókat használó rendszerek megdöbbentő eredményeket produkálnak az olyan mintafelismerési feladatok széles körében, mint például a kézírás, az emberi arcok, hitelkártya-megterhelésekhez hasonló kereskedelmi csalások, és sok egyéb dolog felismerésében.

A *genetikai algoritmusok* (GA) a természet által ihletett másik önszerveződő paradigmához tartoznak, amelyek lemásolják az evolúciót, még a szexuális úton való szaporodást és a mutációkat is. Először is meghatározzuk, mily módon tudunk megfelelő megoldást adni a problémára. Ha a probléma egy repülőgépmotor tervparamétereinek optimalizálása, meghatározzuk a paraméterlistát (minden paraméterhez megfelelő számú bitet rendelve). Ezt a listát tekintjük a genetikai algoritmus genetikai kódjának. Majd véletlenszerűen több ezer vagy még több genetikai kódot generálunk. Minden ilyen genetikai kódot (amely egy sor tervparamétert határoz meg) egy szimulált “megoldási” organizmusnak tekintünk. Ekkor minden szimulált organizmust kiértékelünk egy szimulált környezetben egy meghatározott módszer használatával. Ez az értékelés a genetikai algoritmus sikerének a kulcsa. A példánkban minden megoldási organizmust egy repülőgépmotor-szimulációra alkalmaznánk, és meghatároznánk, milyen sikeres az adott paraméterkészlet a számunkra érdekes kritériumok szempontjából (üzemanyag fogyasztás, végsebesség és így tovább). A legjobb megoldó organizmusos (a legjobb tervek) fennmaradnak, és a többi megsemmisül. Ezután minden fennmaradó organizmust addig szaporítunk, amíg újra el nem érünk egy meghatározott példányszámot.

A GA kulcsa az, hogy az emberi tervező nem közvetlenül programozza a megoldást, hanem hagyja, hogy az spontán módon jöjjön létre a szimulált verseny és javítás ismétlődő folyamatában. Mint már tárgyaltuk, a biológiai evolúció okos, de lassú folyamat, így intelligenciájának felerősítéséhez megtartjuk éleslátását, míg nagyban felgyorsítjuk az ütemét. A számítógép elég gyors ahhoz, hogy sok generációnyi változást szimuláljon órák, napok vagy hetek leforgása alatt. A neurális hálókhoz hasonlóan a GA-k útján hasznot húzhatunk a kaotikus adatokban rejlő finom, ám mélyreható mintákból. A genetikai algoritmusok a káosz- és komplexitáselmélet részeként, egyre inkább alkalmasak egyébként követhetetlen üzleti problémák megoldására, például komplex beszállítói láncok optimalizálására. Úgy tudjuk, a 2001. szeptember 11-i terrortámadásokat a Nemzetbiztonsági Hivatal mesterséges intelligencia-alapú Echelon rendszere előre jelezte, amely a hivatal masszív kommunikációs tevékenységet figyelő rendszerének eredményeit elemzi. Sajnos, az Echelon figyelmeztetéseit az emberi felelősök nem vették figyelembe, amíg már túl késő volt.

(Természettudományok, matematika) A Walesi Egyetem egy eredeti elméleteket megfogalmazni képes mesterséges intelligencia-alapú rendszert ötvöző “robottudóst” fejlesztett ki, amely képes automatikusan kísérleteket végrehajtani, és egy okfejtő egységgel is rendelkezik, amely kiértékeli az eredményeket. A robottudós által kitalált kísérletek háromszor olcsóbbak lettek, mint az emberi tudósok által megalkotottak. A gépet egy csoport emberi tudós ellen tesztelve bebizonyosodott, hogy a gép felfedezései egy szinten álltak az emberekével.

(Fordítórendszerek) A számítógépes fordítás fokozatosan fejlődik, Franz Josef Och, a Dél-Kaliforniai Egyetem számítógéptudósa kifejlesztett egy technikát, amely képes létrehozni egy új nyelvi fordítórendszert bármilyen két nyelv között órák vagy napok alatt, mindössze egy “rosette”-i köre van szüksége – egy szövegre az egyik nyelven és a fordításra a másikon –, bár az igaz, hogy ebből a kettőből több millió szóra. Egy önszerveződő technika használatával a rendszer képes megalkotni a saját statisztikai modelljeit arról, hogyan lett lefordítva a szöveg az egyik nyelvről a másikra, és mindkét irányban továbbfejleszteni ezeket a modelleket.

(Hogyan alkotjuk meg az erős mesterséges intelligenciát?) A 2020-as évek végére részletes modelljeink és szimulációink lesznek az emberi agy minden területéről. Egészen mostanáig az agyba való bekémlelés eszközei nem rendelkeztek térbeli és időbeli dimenzióval, sáv szélességgel stb. ahhoz, hogy az eléggé részletes modellek megalkotásához megfelelő adatokat gyűjtsenek. A 2020-as évek végére képesek leszünk leképező és érzékelő nanorobotokat küldeni az agy kapillárisaiba, hogy belülről is letapogassuk. Az erős mesterséges intelligencia forгатókönyvének egyik egyszerű állítása az, hogy az emberi intelligencia működésének alapelveit az agyi területek visszafejtéses megismeréséből fogjuk kinyerni, és hogy ezeket az alapelveket a biológiai emberi agyénak megfelelő szintű működésre képes számítógépes platformokon fogjuk alkalmazni. Már létezik hatékony eszközkészletünk a gyenge mesterséges intelligenciához. Ezen módszerek folyamatos finomításával, az új algoritmusok kidolgozásával és a többféle módszer bonyolult architektúrákhoz való alkalmazásával a gyenge mesterséges intelligencia nem is lesz olyan gyenge.

(Az erős mesterséges intelligencia hatásai 433-564 p.) “A szingularitás közeledtével ismét fontolóra kell vennünk az emberi élet milyenségére vonatkozó elképzeléseinket, és újra kell tervezni intézményeinket. Például a “G”, “N” és “R” egymásba kapcsolódó forradalmi emberi testünk 1.0-s változatát sokkal tartósabb, többre képes 2.0-s hasonmásokká alakítják át. Nanorobotok milliárdjai fognak mozogni testünk és agyunk véráramában. A testben megsemmisítik a kórokozókat, kijavítják a DNS-hibákat, eltávolítják a káros anyagokat, egészség javítását célzó feladatokat sokaságát végzik el, és mindezek következtében meghatározatlanul hosszú ideig élhetünk öregedés nélkül. (...) A tanulás először online környezetbe tevődik át, de mihelyt agyunk is online lesz, új ismereteket és készségeket is le tudunk tölteni. A munka szerepe - a zenétől és a művészetektől a matematikáig és a tudományokig - mindenféle tudás létrehozása lesz. A játék ismereteket teremt, így nem lesz nyilvánvaló különbség a munka és játék között.

(...)

Az emésztőrendszer áttervezése: a tápanyagokat speciális metabolikus nanorobotok fogják közvetlenül a véráramba juttatni, miközben a vérben és testünkben lévő érzékelők vezeték nélkül kommunikálva dinamikus információt szolgáltatnak a

mindenkor szükséges táplálékról. Az egyik forgatókönyv szerint övünkön vagy a trikónkon fogunk a testbe bőrön vagy más testüregen keresztül bejutó, táplálékot szállító, nanorobotokkal feltöltött speciális tápanyagkészüléket hordani. De az is fontos, hogy a testen belül tartsunk fenn bőséges tartalékokat az összes szükséges anyagból. A test 1.0-s változata nagyon korlátozottan teszi ezt, például a vérben csak néhány percre elég oxigént tárol, míg a glikogénben és más tartalékokban levő kalória, energia mindössze pár napra elegendő. A 2.0-s változat lényegesen nagyobb tartalékokkal rendelkezik, biztosítva, hogy a metabolikus forrásoktól sokkal hosszabb ideig is el lehessünk választva. Vérünk az egyik olyan átható rendszer, amely már tárgya a visszafejtésen alapuló átfogó konceptuális újratervezésnek. Korábban említettem Rob Freitas vörösvértestjeink, vérlemezkéink és fehérvértestjeink helyettesítését célzó nanotechnológiai-alapú terveit. Respirocitái (robotikus vörösvértestjei) lehetővé teszik, hogy órákon keresztül életben maradjunk oxigénbevitel nélkül. Freitas tervei között találunk saját mobilitásukról gondoskodó, önmeghajtó nanorobotikus vértesteket is. Ha a vér önállóan mozog, kiküszöbölhetők a központi pumpálásához szükséges rendkívüli nyomás következtében fellépő mérnöki jellegű problémák is (így a szív feleslegessé válik a test 2.0-s változatában.). A tüdőt az oxigénellátást jócskán feljavító respirocitákkal, oxigént szolgáltató és széndioxidot eltüntető nanorobotokat használva tudjuk kiváltani.

(...)

(*Akkor mi marad?*) Tétélezzük fel, hogy a korai 2030-as években vagyunk, eltávolítottuk a szívet, tüdőt, a vörös- és fehérvérsejteket, vérlemezkéket, hasnyálmirigyet, pajzsmirigyet és az összes hormontermelő szervet, vesét, húgyhólyagot, májat, alsó nyelőcsövet, gyomrot, vékony-és vastagbelet. A csontváz, a bőr, a nemi szervek, az érzékszervek, felső nyelőcső és az agy marad. Nanorobotok összekapcsolásával egy napon azonban lehetővé válik a csontváz fokozatos, beavatkozás nélküli feljavítása, végső fokon kicserélése. Az emberi test 2.0-s változata erős, szilárd és önjavító lesz.

(...)

(*Az emberi agy áttervezése*) Tudósok kísérleteznek “kvantumpöttyökkel” - fotoinduktív (fényre reagáló) félvezető anyagokat tartalmazó parányi chippekkel – is, amelyeket olyan peptidekkel vonnak be, amelyek képesek az idegsejtek felszínének meghatározott pontjaihoz kötődni. Ezzel lehetővé válik, hogy a kutatók az adott hullámhosszú fény segítségével (például gyógyszeradagolásnál) távolról hozzanak működésbe egyes idegsejteket, kiiktatva a beavatkozást igénylő külső elektródákat. Az emberi test 2.0-s változatának forgatókönyve egy régi trend folytatása, mely során egyre bensőségesebb kapcsolatba kerülünk a technológiánkkal. A számítógépek kezdetben távoli, hatalmas gépek voltak, amelyeket légkondicionált termekben őriztek, és kezeltek fehér köpenyes technikusok. Aztán az asztalunkra, majd a karunk alá költöztek, most meg a zsebünkben vannak. Hamarosan rutinszerűen berakjuk őket a testünkbe, az agyunkba. 2030-ra inkább leszünk nem biológiaiak, mint biológiaiak. Mint ahogy a harmadik fejezetben tárgyaltam, 2040-re a nem biológiai intelligencia többmilliárdszor nagyobb lesz a biológiaiainál.

(...)

(*Az emberi test 3.0 -s verziója*) Úgy képzelem, az emberi test 3.0-t - a 2030-as és a 2040-es években – sokkal radikálisabban alakítjuk majd ki. Minden egyes alrendszer újrafarmálása helyett - gondolkodásunk biológiai és nem biológiai részének együttes alkalmazásával – lehetőségünk lesz átalakítani a testünket a 2.0-s változat tapasztalatai alapján. Az 1.0-ból a 2.0-ba való átmenethez hasonlóan a 3.0-ba való átmenet is fokozatos lesz, sok egymással versengő elképzeléssel.

A 3.0-s változat egyik tulajdonságaként a test megváltoztathatóságát képelem el. Ezt virtuálisvalóság-környezetben nagyon könnyen meg tudják tenni, de azok az eszközök is rendelkezésünkre fognak állni, amelyekkel tényleges valóságban is kivitelezhetjük. Az MNT-alapú (molekuláris nanotechnológia-alapú) gyártás magunkba integrálásával gyorsan, tetszés szerint módosíthatjuk fizikai megjelenésünket. Tehát az emberi test 3.0 a mai szabványok szerint emberinek fog tűnni, viszont mivel sokkal formálhatóbb lesz, az idő előrehaladtával kitágul majd a szépségről alkotott elképzelésünk. (...) J. Storrs Hall olyan “fogleteknek” (kb. önszerveződő nanobotok felhőjének az angol *fog* (köd) szó nyomán) elnevezett nanorobotokról írt, amelyek egymással összekapcsolódva sokféle struktúrát képesek létrehozni, és gyorsan át tudják alakítani ezeknek a struktúráknak a felépítését. Azért nevezi *fogleteknek*, mert ha kellő sűrűségben vannak jelen egy területen, képesek különféle hang- és vizuális hatásokat létrehozni a hang és a fény manipulálásával. Lényegében külső (azaz fizikai világbeli) nem pedig belső (idegrendszerbeli) virtuálisvalóság-környezeteket hoznak létre. Egy személy képes módosítani velük a testét vagy a környezetét - habár mivel a *fogletek* a hangot és a képeket kontrollálják, e változások egyike-másika valójában csak illúzió. Hall fogletjei konceptuális szinten olyan valódi, képlékeny testeket képesek létrehozni, melyek felveszik a versenyt a virtuális valóságbeli testekkel.

A 2030-as forgatókönyv szerint a nanotechnológia teljes belemerülést kínál. tökéletesen meggyőző virtuális valóságról fog gondoskodni. A nanorobotok az érzékszervi idegsejtjeink közötti kapcsolat fizikai közelségébe pozícionálják magukat. Ha a tényleges valóságban akarunk élményeket szerezni, a nanorobotok egyszerűen ott maradnak, ahol vannak (a hajszálerekben) és nem csinálnak semmit. Ha be akarunk lépni a virtuális valóságba, elfojtják és a virtuális környezethez illő jelekkel helyettesítik a valódi érzékszerveink felől érkező összes inputot. Ugyanis a test felől jövő - másodpercenként több száz megabitnyi – input érintésre, hőmérsékletre, savszintekre, a táplálék útjára és más fizikai eseményekre vonatkozó információt ábrázol, áthalad a Lamina I. gerincvelői terület idegsejtjein, majd átmegy a hátsó ventromediális magon, és az agykéreg két inzularis régiójába érkezik. Ha hibátlan a kódolás - az agy visszafejtésére tett fáradozásainkból tudni fogjuk, hogyan tegyük azzá - az agy valódiként éli meg a mesterséges jeleket. (...) Mivel változtatni tudjuk a megjelenésünket, és sikeresen más személlyé válhatunk, a virtuális valóságban nem korlátozódunk egyetlen személyiséggé. Egyszerre válogathatunk különféle testek között, különféle személyek részére. Szüleink valamilyenek látnak, míg párunk egy másik testünkkel él át élményeket.

(A tudat kitágítása) 2030 körül a nanobotok legfontosabb alkalmazási területe elménk kitágítása lesz biológiai és nem biológiai intelligenciánk egyesítésén keresztül. Jelenlegi agyunk viszonylag változatlan szerkezettel bír. Jóllehet a tanulási folyamat természetes részeként az idegsejtek közötti kapcsolatokhoz és az ingerületátvivő anyagok koncentrációjához mintázatokot adunk, az emberi agy mai teljes kapacitása erősen korlátozott. A masszívan elosztott nanobotokon alapuló agyi beültetések jelentős mértékben kitágítják memóriánkat, illetve mérhetetlenül feljavítják az összes érzékszervi, mintafelismerő és kognitív képességünket.

(...)

A nem biológiai létezés felé haladva egyre gyarapodnak az önmagunk másolatának elkészítésére (tudásukat, képességeinket és személyiségünket megalapozó mintázatok tárolására) alkalmas eszközök, s mindezek következtében meg tudjuk majd szüntetni a legtöbb ismert halálokat. Az agyfeltöltésről a negyedik fejezetben számoltam be. A

közvetlen agyátvitel forgatókönyve az emberi agy (valószínűleg belülről történő) beszkeneléséből, az összes szembetűnő részlet rögzítéséből és az agy állapotának másfajta - feltehetően sokkal hatékonyabb - számítógépes formában történő ismételt megvalósításából áll. Kivitelezhető eljárás lesz, és a legvalószínűbb hogy a 2030-as évek végére valósul meg. Jelenleg emberi hardverünk összeomlásakor életünk szoftvere – személyes “elmefájlunk” – is vele pusztul el. Viszont abban az esetben, ha birtokoljuk (idegrendszerünk többi részével, belső rendszerünkkel és az elmefájlokban lévő más szerkezetekkel együtt) az agynak hívott mintázatokban kifejezett több billiárdnyi információbyte tárolására és újratárolására szükséges eszközöket, nem lesz többé így. Elmefájlunk élettartama ekkor már nem függ egyedi hardverek folyamatos életképességétől (például a biológiai test és agy túlélésétől). A szoftveralapú emberek végül jócskán túllépnek a jelen emberének korlátjain. Elménk más médiumra másolása számos filozófiai kérdést vet fel, például, hogy tényleg én vagyok-e, vagy valaki más, akinek éppen sikerült összes gondolatom és tudásom birtokába jutni. (...) Szellemi szoftverünk korszakokon keresztül a túlélés miatt kérlelhetetlenül össze volt kapcsolva biológiai hardverünkkel. Azzal a képességünkkel, hogy információfeldolgozásunk összes részletét tárolni tudjuk, és újraformálhatjuk, tényleg szétválna egymástól halandóságunk két aspektusa.

(Terjeszkedés a Naprendszeren túl) Milyen hamar terjeszthetjük ki intelligenciánkat a Naprendszeren túlra? Mint már láttuk, ennek az évszázadnak a végére a Földön a jelen lévő nem biológiai intelligencia több billiószor fejlettebb lesz, mint a biológiai, így nem lesz értelme hús-vér embereket küldeni egy hasonló küldetésre. Nem csupán arról van szó, hogy a biológiai emberek robotszondákat küldenek maguk helyett. Az emberi civilizáció addigra talán minden szempontból nem biológiai jellegűvé válik. Elegendő lenne, ha a szondák önszokszorosító nanobotok lennének. Több billió rajokban indíthatnánk útnak őket, aztán ezek a “magok” meggyökerezhetnének más bolygórendszerekben, és megfelelő anyagok - szén és más szükséges elemek - fellelésével sokasodhatnának, illetve másolatokat készíthetnének magukról. Letelepedésük után a nanobot-kolóniák pusztán információátvitel útján megszerezhetnék az intelligenciájuk optimalizálásához szükséges további információkat, ami csupán energia, nem pedig anyagtovábbítást jelent, így fénysebességgel végrehajtható.

(A tudat nyugtalanító kérdése) A 2020-as évek második felére befejezzük az emberi agy visszafejtését, amellyel az emberek komplexitását és kifinomultságát elérő, majd meghaladó - érzelmi intelligenciával is rendelkező - nem biológiai rendszereket teremthetünk. Egy másik forgatókönyv szerint feltölthetnénk egy létező ember mintáit egy megfelelő nem biológiai, gondolkodó szubsztrátumba. A harmadik, talán legérdekesebb forgatókönyv azt jósolja, hogy az emberek fokozatosan, de végérvényesen átalakulnak biológiai lényekből nem biológiai lényekké. Ez már elkezdődött a rokkantságot és betegségeket kiküszöbölő neurális implantátumok megjelenésével. Ezek után az először orvosi és öregedésmegelőző céllal kifejlesztett, véráramba beoltott nanobotok következnek. Később náluk is kifinomultabb nanobotok jelennek meg, amelyek kapcsolatba lépnek biológiai neuronjainkkal, felerősítik az érzékeinket; vitulis valóságot szolgáltatva az idegrendszeren belül, segítve az emlékezetünket és egyéb rutinszerű kognitív feladatokat látva el. Ekkora már kiborgokká válunk, és erről az idegrendszerünkbe épített ugródeszkáról indulva intelligenciánk nem biológiai része exponenciálisan fejleszti majd képességeinket. (...)

Továbbá ezek a nem biológiai lények képesek lesznek meggyőzni más embereket (biológiai, nem biológiai vagy a köztes életformákat) arról, hogy tudatosak.

(...)

Létezik egy másik, a tudattal kapcsolatos kérdés is: a saját identitásunk. Korábban tárgyaltam annak lehetőségét, miként tudnánk feltölteni egy egyén elméjének mintáit - a tudását, a képességeit, a személyiségét, az emlékeit - egy másik szubsztrátumba (...) Ki is vagyok én? Tudjuk, hogy legtöbb sejtünk hetek leforgása alatt megújul, és még a viszonylag hosszú ideig külön sejteket alkotó neuronjaink is havonta lecserélik minden alkotó molekulájukat. Így tehát teljesen más anyagból vagyok, mint egy hónappal ezelőtt, és csupán az anyag elrendezésének a terve állandó. (...) Bár a másolat mintája ugyanaz lenne, mint az enyém, ám nehéz volna azt állítania, hogy a másolat én vagyok, hiszen még én is léteznék ugyanakkor.

A GNR mélyen összefonódó lehetősége és veszélye (565-619.p.) “Bill Joy (a Java programnyelv egyik kifejlesztője) azt javasolta, hogy korlátozzuk a túl veszélyes technológiákat. Az önreplikáns nanotechnológiát illetően például tiltást javasolt, hasonlóan mint a nanotechnológia két úttörője, Eric Drexler és Christine Peterson. Joy azt is javasolja, hogy ne tegyék közzé a neten a kórokozók génszekvenciáját, és az is szeretné, ha a tudósok önként alkalmaznának bizonyos fokú szabályozását, mert ha megvárjuk a katasztrófát, akkor azután már sokkal súlyosabb és nagyobb kárt okozó szabályozást kell majd elfogadniuk. (...) A veszélyes replikátorok könnyen túl erősek, parányiak és gyorsan terjedőek lehetnek ahhoz, hogy megállítsuk őket – legalábbis, ha nem készülünk fel. (Eric Drexler)

(...)

Az egyik nagy figyelmet kapott valós veszély a nanobotok korlátlan replikációja. Ahhoz, hogy a nanobot-technológia hasznos legyen, több billió ilyen apró, intelligens eszközre van szükség. Ennek a mennyiségnek az eléréséhez muszáj lesz képessé tenni őket önmaguk lemásolására, gyakorlatilag ugyanazzal a módszerrel, ahogy a biológiai világban történik (ahogy a megtermékenyített petesejtből több billió sejt lesz az emberi testben). És ahogy a biológiai sejtosztódás is rossz irányt vehet (a rák esetében), ami akár halált is okozhat, a nanobotok önreplikációját vezérlő mechanizmus meghibásodása - az ún. szürkemassza-forgatókönyv - minden fizikai entitásra, biológiára és egyébre veszélyt jelentene.

Az élőlények – köztük az emberek - lennének az elsődleges áldozata az exponenciálisan terjedő nanobottamadásnak. A nanobottervezés a szén használja alapvető építőelemként. A szénmolekulák képesek láncokba, cikkekbe, gyűrűkbe, nanocsövekbe, lapokba, gömbökbe és számos egyéb alakzatba rendeződni. Mivel a biológia ugyanígy használja fel a szenet, a hibásan működő nanobotok ideális alapanyagforrásnak fogják látni a Föld biomasszáját. Meddig tartana egy elszabadult, önreplikáló nanobotnak elpusztítani a Föld biomasszáját? Egy darab önreplikáló nanobotban józan becslés szerint körülbelül 10 a hatodikon szénatom található. Ennek a rosszindulatú nanobotnak tehát 10 a harminckilencediken számú másolatot kellene készítenie magáról ahhoz, hogy feleméssze a biomasszát. Ezt pedig 130 replikálással érheti el (melyek mindegyike megkétszerezi az elpusztított biomassza mennyiségét). Rob Freitas becslése szerint egy replikáció minimális időtartama körülbelül száz másodperc lenne, tehát 130 replikációs ciklushoz mintegy három és fél óra kellene. Ugyanakkor a pusztítás valódi sebessége kisebb lenne, mivel a biomassza nem hatékonyan van tárolva. A korlátozó tényező a pusztítás frontjának mozgása lenne. A

nanaobotok kis méretük miatt nem képesek túl nagy sebességre. Valószínűleg hetekig is eltartana, míg körbeérnének a bolygón.

(...)

A jövőbeli MNT-gyártóeszközöket meg lehet alkotni olyan biztonsági intézkedésekkel felvértezve, amelyek meggátolnák az önreplikáns nanoeszközök létrehozását. De bizonyos keretek között azért szükség lesz az önreplikálásra. Az önreplikálás képességére ahhoz is szükség van, hogy a nanobotok gyorsan kiterjessék az intelligenciát a Földön túlra, ahogy arról a hatodik fejezetben írtam. A nemkívánatos önreplikálás elleni óvintézkedéseket - például a "szétszórt architektúrát" - azonban az elszánt ellenség, a terroristák kijátszhatják. Freitas számos katasztrofális nanobot-forgatókönyvet is leírt. Az általa "szürke-plankton" forgatókönyvnek nevezett esetben a nanobotok a víz alatt, CH₄ (metán) formájában tárolt, illetve a tengervízben feloldva található CO₂ formában lévő szenet használnának fel. Az óceánban körülbelül tízszer annyi erőforrást találnának, mint a Föld biomasszájában. A "szürke zuzmó" forgatókönyvben a sziklákból és a kövekből nyernék ki a szenet.

(...)

Az is lehetőség, hogy halálos mérgeket előállító géneket adnak hozzá olyan könnyen terjedő vírusokhoz, mint a közönséges nátha és influenza, és ez újabb egzisztenciális fenyegetést jelent. (...) A genetikailag módosított, rosszindulatú vírusok ablaka a 2020-a években fog bezáródni, amikor rendelkezésünkre áll majd a nanobot-alapú, teljes mértékben hatékony antivírus-technológia. Ám az önreplikálódó nanaobotok pedig sokkal nagyobb veszélyt jelentenek majd, mint a biológiai alapúak. A rosszindulatú nanobotok ablakát végül az erős mesterséges intelligencia fogja bezárni, de nem meglepő módon a "nem barátságos" MI még nagyobb egzisztenciális veszélyt fog jelenteni, ahogy arra később kitérek.

(Minél kisebb a kölcsönhatás, annál nagyobb a robbanási potenciál) A közelmúltban vita alakult ki a jövőbeli, nagy energiájú részecskegyorsítók azon képessége kapcsán, hogy szubatomi szinten transzformált energiaállapotok láncreakcióját indíthatják el. Ennek eredményeként a folyamat egyre növekedve széttépne minden atomot a galaktikus környezetünkben. Több ilyen forgatókönyv is született, köztük azzal a lehetőséggel, hogy létrehozzunk egy fekete lyukat, amely elnyeli a naprendszerünket. Ezeknek a forgatókönyveknek az elemzése azt mutatta, hogy igen kicsi ennek a valószínűsége, bár nem minden fizikus ilyen derűlátó a veszélyeket illetően. Gondoljunk arra, hogy valóban egyre kisebb méretskálán egyre nagyobb energiájú jelenségeket találtunk. Alfred Nobel molekulák kémiai kölcsönhatását próbálgatva fedezte fel a dinamitot. Az atombomba, amely több tízezerszer erősebb a dinamitnál, nagy atomok nukleáris kölcsönhatásán alapul. Az atombombánál több tízezerszer erősebb hidrogénbomba még kisebb léptékben működik: kis atomokkal. Noha ebből a felismerésből nem feltétlenül következik, hogy léteznek még erősebb destruktív láncreakciók a szubatomi részecskék szintjén, legalábbis elképzelhetővé teszik a következtetést.

(A lemondás eszméje) Úgy gondolom, hogy megfelelő szintű lemondásnak helyet kell kapnia a XXI. századi technológiák veszélyeire adott etikai válaszunkban. Így a nanotechnológusok lemondanak azoknak a fizikai entitásoknak a kifejlesztéséről, amelyek a természeti környezetben képesek az önreplikációra. Véleményem szerint két kivétel létezik e szabály alól. Előbb-utóbb létre kell hoznunk a nanotechnológia-alapú bolygósintű immunrendszert (a természeti környezetbe beágyazott nanobotokat, amelyek megvédnek bennünket az elszabadult önreplikáló

nanobotokból. A másik kivétel a naprendszerünkön kívüli bolygórendszerek felderítésére bevetett, önreplikáló nanobotokon alapuló szondák iránti igény.

A hasznos etikai vezérelvek egy másik kitűnő példája azoknak az önreplikáló entitásoknak a betiltása, amelyek tartalmazzák az önreplikáláshoz szükséges saját kódjukat. A Ralph Merkle nanotechnológus által “szétszórt architektúrának” nevezett rendszerben az ilyen entitásoknak egy központi biztonságos szerverről kellene beszerezniük ezt a kódot, ami elejét vehetné a nemkívánatos replikációknak. A szétszórt architektúra lehetetlen a biológiai világban, tehát van legalább egy mód arra, hogy a nanotechnológiát biztonságosabbá tegyünk, mint a biotechnológiát. Más szempontból a nanotechnológia sokkal veszélyesebb, mivel a nanobotok fizikailag erősebbek lehetnek, mint a fehérjealapú entítások.

(...)

További óvintézkedések: A replikáció olyan alapanyagokat igényeljen, amelyek nem találhatóak a természetben; a gyártóeszközök legyenek képesek a végtermékek előállítására, de önmaguk ne tudjanak replikálódni, és a végtermékek se legyenek képesek erre. A replikációhoz legyen szükség titkosított és időben korlátozott érvényességű replikációs kódokra. A korábban említett szétszórt (drót nélküli utasításokon alapuló) architektúra egy példa erre.

(Védelem a “barátságatlan” erős MI-vel szemben) Ám akármilyen hatékony mechanizmus is a szétszórt architektúra, az erős MI-vel szemben nem fog megvédeni. A szétszórt architektúra által biztosított gátak a nanoentítások intelligenciájának hiányán alapulnak. Az intelligens entítások azonban elég okosak ahhoz, hogy megkerüljék ezeket a gátakat. Eliezer Yudkowski számos paradigmát, architektúrát és etikai szabályt elemzett, melyek segítenek biztosítani azt, hogy amikor az erős MI rendelkezésére állnak az eszközök, amelyekkel elérheti és módosíthatja a saját felépítését, akkor is barátságos maradjon a biológiai emberiséghez, és támogassa annak értékeit. Tekintettel arra, hogy az MI-t nem lehet visszahívni, Yudkowski hangsúlyozza, hogy “elsőre jól kell megcsinálni”, és már a megszületését célzó tervnek “nulla javíthatatlan hibája lehet”. De eredendően nem lesz abszolút védelem az erős MI-vel szemben.

*(A gazdag-szegény választóvonal kritikája)*Jaron Lanier és mások felemlegetnek egy további “vérfagyasztó” lehetőséget, miszerint az új technológiák olyan előnyökhöz és lehetőségekhez juttatják a tehetőseket, amilyenekhez az emberiség többi részének nem lesz hozzáférése. A társadalmi különbségek persze mindig léteztek, de a gyorsuló megtérülések törvénye fontos és jótékony változást hoz majd ezen a területen is. Az ár-teljesítmény arány folytonos exponenciális javulásával ezek a technológiák hamar olcsóvá, majd hógynem ingyenessé válnak.

I.2. Nick Bostrom: “Szuperintelligencia”

Bostrom könyve 2014-ben jelent meg az eredeti angol nyelven, és egy évvel később magyar fordítása is. Így a 2005-ben íródott Kurzweil könyvhöz képest majd egy évtizeddel több tapasztalatot tudhat e könyv a háta mögött a mesterséges intelligencia fejlődése terén. Annyiban is eltér ez a könyv a másiktól, hogy a mesterséges intelligencia a létrejöttének és működésének különféle formáit tisztábban szétválasztja és külön-külön is megnézi ezek létrejöttének esélyeit, illetve előnyeit és veszélyeit. Itt is azt az ismertetési módot választom, hogy a fontosabb állításokat, és

téziseket tartalmazó szöveget kivonatolom, és a saját értékelést ezekről az olvasónaplót követő részben helyezem el.

(A szuperintelligenciához vezető utak) A szuperintelligencia olyan intellektus, amely szinte minden fontos területen jóval túlesz az ember képességein. Egy lehetőség a mesterséges intelligencia megvalósítására, ha veszünk egy agyat, és egyfajta mintadarabként használjuk a gépi intelligencia létrehozásához. Korábban már említettük az emberi agy által inspirált MI-technológiát, a neurális hálózatokat. A gépi intelligencián dolgozó szakemberek különbözőképpen vélekednek az idegtudományi megközelítés előnyeiről a teljesen mesterséges rendszerekkel szemben. Turing gondolata, hogy olyan programokat fejlesszünk ki, amely saját tartalma nagy részét tanulásból nyeri, nem pedig előzetes programozásból, mind az idegtudományi, mind a mesterséges megközelítés szempontjából érdekes lehet.

Turing gyermekgépének létezik egy változata, az “MI-mag” gondolata. Turing által elképzelt gyermekgép relatíve fix architektúrával rendelkezne, amely a tartalom felhasználásával fejlesztené belső képességeit. Az MI-mag ezzel szemben kifinomultabb mesterséges intelligencia lenne, amely képes fejleszteni saját architektúráját is. Ezzel eljutottunk egy újabb fontos fogalomig, a “rekurzív önfejlesztésig”. A sikeres MI-mag képes lenne iteratív módon fejleszteni önmagát: az MI korai verziója meg tudná tervezni saját fejlettebb változatát, ez utóbbi pedig - mivel okosabb lenne, mint az első - még okosabb változatot tudna tervezni, és így tovább.

(A teljes agy emulációja) A teljes agy emulációjánál (azaz “feltöltésénél”) az intelligens programokat egy biológiai agy beszkenelésével és a struktúrájának pontos modellezésével alkotnánk meg. A teljes agy emulációjához a következő lépésekre lenne szükség. Először is részletesen be kell szkennelni egy emberi agyat. Ez történhet úgy, hogy a halál után vitrifikációval stabilizálják az agyat (az eljárás egyfajta üveggé változtatja a szövetet). Így vékony szeletekre lehetne vágni a szöveteket, amelyek beszkenelhetők lennének, például elektormikroszkópok segítségével. Ebben a szakaszban több szkennert is dolgozhatna párhuzamosan, hogy egyszerre szkenneljenek be több agyszeletet. Második lépésben a szkennerek nyers adatait kellene betölteni egy számítógépbe automatikus képfeldolgozásra, hogy megkapjuk az agy háromdimenziós neurális hálózatát, amelyben a gondolkodás végbemegy. Az elkészült térképet aztán egyesíteni lehet a különböző idegsejtekre vonatkozó neuroszámítási modellekkel (mint például a különböző szinoptikus kapcsolatok). A harmadik fázisban az előző lépés eredményeként létrejött neuroszámítás-technikai modelleket lefuttatjuk egy megfelelő kapacitású számítógépen. Ha a folyamat maradéktalanul sikeres, akkor az eredmény az eredeti tudat teljes leképezése lenne, emlékekkel és személyiséggel együtt. Az emulált emberi agy onnantól számítógépen létezne tovább. Ez az elme vagy a virtuális valóságban lakna, vagy robotikai mozgatószervekkel érintkezne a külvilággal.

A teljes agy emulációjához nem kell rájönnünk, hogyan működik az emberi értelem, vagy hogyan programozunk le egy mesterséges intelligenciát. Csak arra van szükségünk, hogy megértsük az agy alapvető számítási elemeinek alacsony szintű funkcionális tulajdonságait. A teljes agy emulációjához nem szükséges semmilyen alapvető fogalmi vagy elméleti áttörés. (...) Egy részletesebb elemzés megkülönböztetné az emuláció sikerességének különböző fokozatait az alapján, hogy az emulált agy milyen információ-feldolgozó képességet őrzött meg az eredetiből. A következő kategóriákat állíthatnánk fel: 1. *nagy pontosságú emuláció*, amely rendelkezik az emulált agy összes tudásával, képességeivel és értékeivel; 2. *torz*

emuláció, amelynek működése egyes területeken jelentősen eltér az emberitől, de nagyjából képes az emberi agy munkájának elvégzésére; 3. *generikus emuláció* (amely szintén torzíthat), amely inkább hasonlít egy gyerekhez. Nincsenek meg azok a képességei, vagy emlékei, amely egy felnőtt agy emulációjának meglennének, de képes lenne megtanulni nagyjából mindent, amit egy ember. (...) (*Melyik fog először megvalósulni?*) Az MI-forgatókönyhöz hasonlóan nagyobb biztonsággal kijelenthetjük, hogy az emulációs út nem a közeli jövőben hoz majd eredményt (például a következő tizenöt évben) mert tudjuk, hogy még az elinduláshoz szükséges technológiák kifejlesztése is várat magára. Ezzel szemben úgy tűnik, hogy egy MI-magot lehet kódolni a mai számítógépeken is, és elképzelhető – bár nem valószínű -, hogy valaki valahol már rendelkezik a tudással, hogy ezt a közeli jövőben megvalósítsa.²

(*Biológiai gondolkodás*) Az emberinél nagyobb intelligencia kifejlesztésének harmadik útja az, ha a létező emberi agyakat használjuk fel, és megnöveljük azok képességeit. (...) Az őssejtekből származtatott ivarsejtekkel az egy pár számára rendelkezésre álló kiválasztási lehetőség nagyban megnövekszik. Az őssejtekből származtatott ivarsejt esetében néhány sejtől gyakorlatilag végtelen számú ivarsejt teremthető, amelyek kombinációjából embriók hozhatók létre, amelyek ezután genotipizálhatók vagy szekvenálhatók, és a legígéretesebb kiválasztható beültetésre. Ezzel a módszerrel tíz vagy még annál is több generációnyi szelekciót valósíthatnánk meg csupán néhány év alatt. Az ilyen úton fogant egyének átlagos intelligenciája nagyon magas lenne, a teljes emberi populáció legintelligensebb egyedével egyenlő, vagy még magasabb. Egy olyan világ, amelyben a népesség minden tagja ilyen módszerrel született, kollektív szuperintelligenciaként működhetne. (...) A folyamat terjedésével változhatnak a szülői normák is, különösen a társadalmi elit körében. Az új normák szerint a felvilágosult, felelősségteljes pároknak előnyben kell részesíteniük a kiválasztást a természetes fogantatással szemben. Akik kezdetben vonakodtak, később szintén a szelekció mellett dönthetnek, hogy gyerekük ne kerüljön hátrányba a barátaik és kollégáik tehetségesebb gyerekeikkel szemben.

(...)

Az igazi szuperintelligenciát (ellentétben a jelenlegi intelligencia szintjének kis növelésével) először valószínűleg az MI útján fogjuk elérni, de a teljes agy emulációja is bizonyulhat a szuperintelligenciához vezető leggyorsabb útnak. Mivel ez a forgatókönyv főleg a meglévő technológiák fokozatos fejlesztését igényli, nem pedig elméleti áttöréseket, nagy rá az esély, hogy valamikor megvalósul. Azonban valószínűnek tűnik, hogy ha a teljes agy emulációja útján gyorsan is haladhatunk, a mesterséges intelligencia fogja átszakítani a célvonalat, mivel lehetőség van részleges emulációkra alapozott neuromorf MI-k kifejlesztésére. (...) Az értelmi képességek biológiai növelése nyilvánvalóan járható út, főleg a genetikai kiválasztás segítségével. A sorozatos embriószelekció jelenleg nagyon ígéretes területnek tűnik. De így csak legfeljebb gyenge szuperintelligenciák alakulhatnak ki.

(...)

(*A szuperintelligencia formái*) Ezek a következők: a gyors szuperintelligencia, a kollektív szuperintelligencia és a minőségi szuperintelligencia, de csak a legutóbbi az, ami az igazán erős mesterséges intelligencia körébe tartozik.

² Az e könyvvel párhuzamosan írt agykutatási könyvében Kurzweil azt írta, hogy az agy emulációjával foglalkozó svéd idegtudós, Henry Makram 2009-ben azt mondta neki egy konferencián az előrehaladó sikerei kapcsán, hogy 2023-ra kész lesz az emberi agy teljes emulációjával.

(Egy intelligenciारobbanás dinamikája) Lényeg: attól a ponttól, hogy a gépek emberihez hasonló szintre érnek az általános értelmi képességek terén, vajon mennyi idő lesz, amíg eljutunk az emberinél lényegesen erősebb szuperintelligenciához?

Három fejlődési forgatókönyvet vázolhatunk:

- Lassú átmenet, mely évtizedeket vagy évszázadokat vesz igénybe.
- Gyors átmenet, mely igen rövid idő alatt, percek, órák napok alatt hozná létre az intelligenciारobbanást.
- Végül a közepes átmenet, amely hónapok vagy évek alatt tenné ezt.

Megpróbáljuk megindokolni, hogy miért valószínűtlen hogy lassú átalakulás forgatókönyve. Ha és amikor az átmenet megindul, nagy valószínűséggel robbanásszerű lesz. Két fázist különböztethetünk meg. Az első az átmenet elején kezdődik, amikor a rendszer egyéni intelligenciája eléri az emberi szintet. Ha a rendszer olyan kiterjedt képességekre tesz szert, hogy az őt fejlesztő optimalizációs eredmény nagyobb részét már saját maga szolgáltatja, egy második növekedési fázis kezdődik meg. Ez alapvetően változtatja meg a dinamikáját, mert a rendszer teljesítményének bármekkora növekedése egyúttal a saját maga fejlesztésére fordított optimalizációs erőt is arányosan növeli. A duplázási konstans az adott forgatókönyvtől függ, de akár nagyon alacsony is lehet – egyes forgatókönyvek szerint csak néhány másodperc.

(Döntő stratégiai fölény) A dinamikától némileg elkülönülő, ám ahhoz kapcsolódó kérdés, hogy ha létrejön a szuperintelligencia, akkor vajon egy ilyen elme lesz, vagy több? Elképzelhető, hogy egy projekt intelligenciारobbanás útján annyival fejlettebbé válik, hogy onnantól az a projekt diktálja a jövőt? Ha az átmenet gyors (néhány óra, nap vagy hét) akkor valószínűtlen, hogy két egymástól független projekt egyszerre induljon fejlődésnek; szinte biztos, hogy mire a második projekt elindulna, az első már a fejlődési görbe felső ágára ér. Ha az átmenet lassú (évek, vagy évtizedek alatt játszódna le) akkor több projekt is fejlődésnek indulhat, és bár az átmenet végére mindegyik hatalmasat fejlődne, nem lenne olyan időszak, amelyben az egyik projekt elég nagy előnyre tenne szert a döntő fölény kiharcolásához. Ha pedig egy projekt-döntő stratégiai fölényt harcol ki, akkor kérdés, hogy a vetélytársakat elnyomva egyeduralkodói pozícióra tör-e. Tartsuk észben, hogy az egyeduralom elvontan értelmezett fogalom: lehet demokrácia, diktatúra, egyetlen domináns MI, a világ normáiból alkotott erős rendszer, amelyben a végrehajtó hatalom is benne van, vagy akár egy gépi hűbérúr - lényeg, hogy az egyeduralkodó olyan szereplő lenne, amely képes megoldani minden, a világ irányításában felmerülő koordinációs problémát.

(...)

Ha a digitális intelligencia egy emberi elme másolata (ahogy egy teljes agy emulációjánál lenne), akkor a digitális másolat örökölné az emberi mintapéldány motivációit. A másolat alapján létrehozott elme még akkor is megtarthatna egy-két motivációt, ha jelentősen módosítanák, hogy szuperintelligenssé fejlesszék.

(Élet egy algoritmikus gazdaságban) Az emberek életmódja az átmenet utáni malthusi társadalomban talán minden történelmi szereptől különbözni fog. Az emberek nagy része egyszerű járadékos lesz, aki megtakarításaiból tartja fenn magát valahogy. Lehet, hogy nagyon szegények lesznek, hiszen csak minimális vagyonukra és némi állami támogatásra számíthatnak. A fejlett technológia világában élnek majd, és nemcsak szuperintelligens gépek között, hanem az öregedésgátló gyógyszerek, virtuális valóság és számos javítótechnológia és örömszerző drog uralta környezetben. Ezek nagy része

azonban megfizethetetlen lesz a számukra. Az öregedésgátló gyógyszerek helyett talán olyan készítményeket szednek majd, melyek lassítják a növekedésüket és az anyagcseréjüket, hogy csökkenjen a megélhetési költségük (a gyors anyagcseréjük nem tudnának megélni az egyre kisebb bevételeikből). Ahogy az emberek száma nő, és az átlagjövedelmük csökken, visszavonulhatnak egy olyan létállapotba, amelyben még jogosultak lesznek a járadékaikra - talán tartályban tárolt, de öntudatuknál lévő agyak formájában, amelyeket gépek tartanak életben; így lassan meg tudnának takarítani annyit, hogy egy robottechnikus megalkossa klónjukat, és így szaporodhassanak. Ha még jobban takarékoskodnak, feltölthetik magukat a számítógépes rendszerbe, egy fizikailag optimalizált tárolóba, amelyet a szuperintelligencia alkotott. Ez a létezési forma valószínűleg hatékonyabb lenne, mint egy biológiai agy. A digitális világba történő migrációt azonban megnehezíthetik, ha az emulációkat nem fogják embernek vagy járadékra jogosultnak tekinteni. Ebben az esetben a biológiai embernek lehet még némi hely, bár az emulációk vagy a mesterséges intelligenciák populációja sokkalta nagyobb lehet.

Megjegyzés: az eddigi a kivonatok Bostrom könyvének 237. oldaláig terjedő részből valók, ezután a könyv másfél száz oldalon azokat a technikákat igyekszik kitalálni és elemezni, amelyek egy ember fölé növe szuperintelligencia ember ellen fordulása esetén védekezést tennének lehetővé. Röviden ezeket úgy lehet összegezni, hogy Bostrom végső soron lehetetlennek tartja, hogy fizikailag el lehessen szigetelni a szuperintelligenciát attól, hogy hozzájusson világ erőforrásaihoz. Másik út így csak az lehet a fizikailag korlátozhatatlan szuperintelligencia emberi társadalommal együttélésére, ha eleve olyan értékekhez, normákhoz és morális elvekhez próbáljuk kötni, amelyek biztosítják, hogy hatalmas technológiai változtatási képessége mellett sem söpri el az emberi létezés életfeltételeit. Ez utóbbiak azonban megítélésem szerint kissé naivak a könyvben, és úgy lehet összegezni ezek elégtelenségét, hogy egy technológiai célokban gondolkodó, önmagát milliószeres gyorsasággal a célok szerint változtató és átépítő szuperintelligencia, az erős MI a belé programozott értékeket és normákat mint korlátokat és gátakat söpörheti el legelőször. Így csak a pesszimizmus maradhat ezeket végig olvasva.

II. A mesterséges intelligencia társadalomváltoztató kilátásai

II.1. A gépi értelem erős MI alakja

A gépi értelem erős mesterséges intelligencia (MI) alakja azt a fokozatot jelzi, amikor a mesterséges intelligencia az emberi értelem szintjét eléri, majd azt ezerszeresen és milliószeresen meghaladja, szemben a gyenge MI ma ismeretes teljesítményével. Előkérdésként még az is felmerül, hogy ez lehetséges-e egyáltalán, és tényleg létre tud-e jönni ilyen teljesítményű mesterséges intelligencia? Ám az elemzések az eddigi exponenciális gyorsasággal növekvő teljesítményéről, ezt könnyen megválaszolhatóvá teszik: igen létrejön, kérdés csak az, hogy ez mikor lesz, már 2040 körül vagy csak a 2100-as évek végéhez közeledve. Ezt szem előtt tartva így két alapkérdés merül fel: 1) elszabadul-e az ilyen fokozatú mesterséges intelligencia az ember és az emberi társadalom intézményi felügyelete és irányítása alól; 2) a második kérdés az, hogy ez az elszabadult mesterséges intelligencia milyen természetű lesz, önálló öntudattal és átható akarattal rendelkezik-e, amely az emberek kívánalmaitól függetlenül irányítja hatalmas valóságalkító kapacitását, vagy ezzel

szemben ez az önálló átlátó akarat nem tud létrejönni párhuzamosan az óriási technikai kapacitással együtt, és csak mint egy naiv kisgyerek átlátási képessége lesz párban az óriási technológiai változtató kapacitással.

Az első kérdést - az elszabadulás lehetőségét - illetően a gépi értelem fejlesztésének fő irányaként látott genetikai algoritmusokon és rekurzív önfejlesztésen alapuló öntanulás és önmegváltoztatás következményeiből kell kiindulni. Ennek révén csak a kezdetekben betáplált paraméterek megválasztásában van meg az emberi ellenőrzés és behatás, de ezután egyrészt az ezek megvalósításra az embertől elszakadó megoldási javaslatok és a gépi értelmet irányító mechanizmusok jönnek létre a gépi értelem belső irányításában, másrészt a betáplált paraméterek módosítása is az öntanulás hatókörébe kerül. Beleértve ebbe még azt is, hogy a rekurzív önfejlesztés technológiájával még maga a megoldásokat hordozó hardver megváltoztatása is létrejön. Ezek már mind ma is léteznek, de a mai gyenge MI szintjén az emberi értelem fölénye az öntanulás és rekurzív önfejlesztés kicsúsztatását az emberi irányítás alól még meg tudja gátolni. Ám a maihoz képest milliószoros gyorsasággal végbemenő öntanulási ciklusok és rekurzív önfejlesztési ciklusok másodpercekre, percekre és órákra csökkenése naponta akár százszoros alapvető változásokat tudnak majd létrehozni, amelyek már túl lesznek az emberi értelem általi ellenőrzés lehetőségén. Az erős MI emberi ellenőrzéstől való elszakadása így egyszerűen következik a mai tendenciákból egy pont után.

A következő kérdés így az, hogy milyen természetű lesz ez az emberi ellenőrzés alól kiszabadult mesterséges intelligencia. Ennek megítélésére egy disztinkciót érdemes tenni, és külön kell választani az intelligencián belül a *technológiai intelligenciát*, és külön az átfogó valóságban és az emberi *társadalom valóságában eligazodás képességét*. A technológiai intelligencia a fizikai-biológiai világ átalakítási képessége, és e képesség még egy külön aspektusának nagysága arra vonatkozik, hogy ez mennyiben megakadályozhatatlan más erők (így az ember) behatása révén. Tehát ez a fizikai-biológiai világ maga alá gyűrésének és uralásának képessége. Ez az, ami növekszik egyre inkább a mesterséges intelligencia területén, míg az átfogó valóság, benne az emberi társadalom valóságának átlátási képessége, benne az emberi társadalom fennmaradását biztosító feltételek átlátása messze elmarad ettől. Nick Bostrom ez utóbbinak az MI-programba beépítés fontosságát elemezte könyvében, és ezek alapján itt két problémára lehet rámutatni. Az egyik probléma az, hogy az emberi társadalom alapszerkezetének és ennek fennmaradásának nincs egyetlen objektív paraméter-rendszere, hanem az egyes társadalmi csoportok értékválasztásaitól függően változik az értékek kiválasztása és hierarchiába állítása. Attól függően, hogy melyik társadalmi csoport dominál, választhatók ki a legkülönbélebb értékek és értékhierarchiák. Ez azonban még a kisebb baj. A nagyobb baj onnan kezdődik, hogy még az így szeletíven és hiányosan - és esetleg a társadalom nagyobb része számára hátrányos értékhierarchia - is a rekurzív önfejlesztés áldozatává válhat az elszabadult erős MI változtatásai révén. Ha önmaga tudja értelmi komponenseit újra és újra megalkotni, akkor semmi biztosíték nincs arra, hogy ne váljanak a technológiai paraméterek akadályává a betáplált társadalmi értékpremisszák, és ne változtassa meg, tüntesse el ezeket a mesterséges intelligencia az önváltoztatási folyamatainak már néhány ciklusa után.

Így pontosabban megadva a disztinkciót, ez a *technológiai intelligencia és a társadalmi értékekre vonatkozó intelligencia szembenállásában* fogható meg. Miközben a mesterséges intelligencia a technológiai intelligencia dimenziójában óriásivá válik, addig a társadalmi értékekhez és ezek ütköztetéséhez, illetve feloldásukhoz szükséges intelligencia dimenziójában a buta kisgyerek szintjén

maradhat. Ha pedig - ezt feljavítandó - az értékekre és ütköztetésükre, feloldásuk kezelésére külön algoritmusokat illesztnek be a mesterséges intelligencia programjába, semmi biztosíték nincs arra, hogy ne kerüljenek kiiktatásra rövid idő után az öntanulás és a rekurzív önfejlesztés révén. Ebből következően megítélésem szerint nem a sokszor olvasható leírás a megfelelő a “gonosszá vált” mesterséges intelligenciáról, hanem az átfogó valóság és az emberi társadalom valóságának átlátásával nem rendelkező, vak-buta és ennek ellenére hatalmas technológiai intelligenciától kell félnünk. Ez a mesterséges intelligencia nem gonoszságból tolhatja félre az emberi társadalom létezési feltételeit, ha óriási technológiai kapacitásával ez meg tudja tenni, hanem egy buta kisgyerek átlátó képességi hiányai miatt. Már itt jelezni kell azonban, hogy az emberi agy emulációjának (az elme feltöltésének) utóbbi években felerősödő fejleményei ez utóbbi problémát módosíthatják (lásd majd a következő részben).

Ezt az összefüggést kihangsúlyozva, az erről gondolkodó kutatók az ember fölé növekvő, hatalmas mesterséges intelligencia társadalomra veszélyességének megfékezésére legalább programjának kezdeti algoritmusában igyekeznek olyan működési elveket elhelyezni, melyek biztosíthatják az emberi társadalom számára veszélyes fordulatokat elkerülni. Egy újonnan megjelent kötetben Ben Goertzel és Joel Pitt közös tanulmányukban az emberi társadalom felé pozitív elfogultságot biztosító programelemek lehetőségeit igyekeztek felmérni a mesterséges intelligencia tervezésénél. Abból indultak ki, hogy ennek biztosítását ugyan nem lehet teljes mértékben garantálni, de hogy legalább ebbe az irányba ösztökéljék működését, azt be lehet iktatni a programjába: “Our current perspective is that provably, or otherwise guarentee-ably, Friendly AI is not achievable. On the face of it, achieving strong certainty about the future behaviors of being massively more generally intelligent than ourselves seems implausible. Again, we are aiming at a more modest goal – to explore ways of biasing to odds, and creating AI systems that are significantly more likely than not to be Friendly.” (Goertz/Pitt 2014:65.) Így a gépi értelem “barátságos” irányának biztosítására az első imperatívusz, amit szükségesnek tartanak beépíteni az egyre hatalmasabbá váló gépi értelem programjába, hogy a rekurzív önfejlesztés ciklusai az első időszakban a lehető leglassabbak legyenek – az emberi értelem felfogóképességére szabva még -, és hogy ez az első ciklusokban még ne mehessen végbe teljesen önállóan, az emberi közreműködés nélkül. Ugyanígy az etikai elvek jól végiggondolt beépítését is javasolják a gépi értelem programjába, és ezek gazdag példatárral ellátását, illetve ezt segítő, az erre ösztökélő szituációkon való sokszoros kipróbálását, végigfuttatását javasolják még a kezdeti fázisokban. “Where, after an AGI (Artificial General Intelligence) has learned some of the everyday aspects of justice, including the balance of justice with empathy in everyday life, and once it has also gotten familiar with the application of abstract ethical principles to other aspects of ordinary life, it will be well poised to appreciate ethical principles and their utility in making difficult decision. It will be able understand the abstract nature of justice in a richer and more holistic way.” (Goertz/Pitt 2014:72.) Végző fokon azonban még így is csak reménykedni lehet, hogy a teljesen önjáróvá vált és hatalmas változtatásokra képes erős mesterséges intelligencia nem számolja fel az emberi társadalom működésének előfeltételeit.

Ezek után felmerülhet az elemzőben, hogy ha ennyire evidens az ezzel foglalkozók számára az erős mesterséges intelligencia potenciálisan az egész emberiséget veszélyeztető vonásainak létrejötte - a jótékony áldásai és előnyei mellett -, akkor miért nem állítják inkább le az egészet, amíg ez lehetséges?! Erre kitűnő elemzések találhatók Bostrom magyarul is megjelent és az előbb ismertetett

könyvében, de Kurzweil is többször utal erre. A mesterséges intelligencia kutatásai ugyanis már évtizedek óta a nagyhatalmak katonai és világhatalmi vetélkedésének egyik területét jelentik, hatalmas pénzösszegeket és kutatási kapacitásokat erre fordítva, így a legelől tartó kutatási eredmények e téren mindig titkosak is. Lényegében ugyanaz történik e téren is, ami az 1940-es évek végétől kezdve az atomkutatások és az atomfegyver kifejlesztése idején történt, és amelyik nagyhatalom e téren előnybe kerül, az a világ egészének uralásában nagy lépést tesz előre. Az egész emberiséget fenyegető kutatásokról lemondani a gépi értelem terén így azt jelentené a legélesebb konkurencia közepette, hogy az ezt megtevő átengedi a versenyelőnyt a vetélytársainak. Ehhez kapcsolódó probléma, hogy még ha egy idő után a növekvő veszély láttán mégis elkezdenének erőfeszítéseket tenni a világ e téren vetélkedő legnagyobb hatalmai, akkor is gondot okoz az, hogy nem lehet előre látni azt a minőségi átbillenő pontot a mesterséges intelligencia fejlesztésében, amikor a mai, még gyenge mesterséges intelligenciának minősülő gépi értelem végleg átbillen az erős és már kontrollálatlanná váló utóda felé. Mind Kurzweil, mind Bostom elemzésében az atomkorszakhoz képest az erős mesterséges intelligencia korszaka ezerszer nagyobb áldásokat és szétágazóbb változtatásokat hoz létre az emberi társadalmak mai állapotához képest - és a robotika növekvő fejleményei ezt már ma is mutatják a mindennapokban. Ám fenyegetései is éppúgy ezerszer nagyobbak, mint azt az atomkorszak létrejöttével is észlelték a korábbi háborús vetélkedésekhez képest. Leállítani azonban éppúgy lehetetlen, mint ahogy az atomért való küzdelem megállítására sem volt lehetőség. Az új MI korszak (az erős mesterséges intelligencia korszaka) csak abban tér el e vetélkedéstől, hogy míg az atomkorszakot pacifikálni lehetett végül azzal, hogy az egymást elrettentő atomhatalmak megegyezhettek egy idő után az atomfegyverek ellenőrzésében és a többiek ettől megtiltásában, addig az erős MI létrehozása az egyik világhatalom részéről egyben az ő háttérbe tolását is maga után vonja, és ő is bekerül az egész emberiséggel együtt a fenyegettek közé. Nem lesz, aki pacifikálja az új domináns erőt, és ezt csak maga tehetné meg.

II.2. Az emulált emberi agy digitális létezése

Az erős mesterséges értelem létrejötte a gépi értelem már működő, gyenge alakjából már több évtizede szem előtt tartott lehetőség, ezzel szemben az "elme-feltöltés" (*mind uploading*) vagy más elnevezésben emberi agyi emuláció csak az utóbbi évtizedben került a fokozódó érdeklődés középpontjába. Ez is a gépi értelem, a mesterséges intelligencia egyik ágát jelenti, de itt az előbbi mellett más ösztönzők adják a fő motívumot. Ugyanis, ha sikerülne a teljes emberi agyi emuláció, és az eddigi biológiai folyamatok helyett komputeres futtatások is lehetővé tennék az agy idegfolyamatainak működését, akkor ennek egyik perspektívája, hogy az ember elméje és személyisége -megszabadulva az elenyészésnek kitett biológiai test általi hordozás kizárólagosságától - egy örökéletű hordozóra is átkerülhetne. A Transzcendens c. film Johny Depp főszereplésével néhány éve ezt dolgozta fel, de a kutatások is a legintenzívebben folynak, és például az elmúlt években az Európai Bizottság is egy másfélmilliárd eurós összeget adott erre a kutatási célra, kezdve a legegyszerűbb élőlények kisméretű agyának emulációjával, majd a kisebb emlősök és párhuzamosan ezzel az emberi agy emulációjának előrehaladásával.

A tiszta gépi értelemhez képest az emberi agy emulálásával létrehozott önálló létezés kérdései a mai technikai állapotok szintjén még kevéssé láthatók át. Még az első, emlős szintű emuláció sem jött eddig létre, és a korábban már jelzett 2014-es

dátum megvalósulása Henry Markram emulációi révén sem kaptak híradást, így ennek csúszását kell feltenni. (Egy új hír a világsajtóban arról szól, hogy 2017-re várható egy patkány agyának teljes emulációja.) Csak ezután lehet megalapozottabban tudományosan elgondolkodni azon, hogy az ilyen agyi emuláció mennyiben ismétli meg az eredeti létezőt, és hogy gondolkodási funkciói hogyan működnek, esetleg mennyiben térnek el a fizikai-biológiai testtel rendelkező létezőétől. Az emulációk reális létezése nélkül ezért tényleg csak a filozófiai szintű kérdések szintjén lehet erről gondolkodni. Persze ez sem haszon nélküli, ám mindenképpen csak spekulatív szintű lehet, és ezt szem előtt kell tartani a következőkben.

Az emberi agy emulációjánál előkérdés még maga a technikai megvalósíthatóság is, az, hogy a sok milliárd agysejt sok billiónyi kapcsolódásaihoz - szinapszissaihoz - szükséges számítógépes kapacitás (tárhely és gyorsaság) létrejöhet-e egyáltalán. Az elemzések az eddigi exponenciális gyorsaságú fejlődést alapul véve a mai elégtelenség után ezt kb. harminc éves fejlődéssel elérhetőnek tekintik, így ezzel nem lesz probléma. Egy új híradás szerint például egy emberi agy neurális - tehát a legrészletesebb működésének - szintjén végbemenő folyamatok egyetlen másodpercnyi hosszát emulálva ("lefejtve") és számítógépes formátummá átalakítva a világ leggyorsabb számítógépén futtatva, negyven percet vett igénybe ez a futtatás. Vagyis ma még 2400 másodperc alatt lehet reprodukálni egyetlen agyi másodperc folyamatait a számítógépeken, és ez első végiggondolásra elbátortalanítóként hathat. Ám ha a Moore-törvény jövőbeni érvényét továbbra is feltesszük, vagyis a számítógépes teljesítményeknek kb. másfél évenként való duplázódását - amire pl. a kvantumszámítógépek gyors előrehaladása is feljogosít -, akkor a 2400-szeres gyorsulást kb. 15-16 év alatt elérjük, és így az emberi agy folyamatai egy az egyben tudnak futni időben már a komputeren is. A fő vita inkább az lehet, hogy az elme összes tartalmának sikeres emulációja, "lefejtése" és számítógépen való megisméltése átviszi-e egyben az eredeti elme öntudatát is, és e tartalom számítógépen való futtatása közben egyben az öntudat is mindig megjelenik-e?

Az igen intenzív gondolkodás és vita e kérdés felett okos disztinkciókat eredményezett az utóbbi években. Külön kell választani ezek szerint azt, hogy az agyi emuláció után a komputeres futtatásban működni fognak-e az egyes mentális folyamatok, és ettől egy külön kérdés, hogy ezek összegződéséként az ilyen emulált emberi agyi működés egyben *létrehoz-e egy egységes öntudatot*, mintegy a párhuzamosan a folyamatosan futó mentális folyamatok *együtt-látójának* a pozícióját?! Végül egy harmadik kérdés, hogy ha igen, akkor ez az öntudat az emulálás előtti, eredeti elme öntudata lesz-e, vagy ez egy új öntudat keletkezését jelenti, amelynek csak annyi köze lesz az eredeti emberi elme hordozójához, hogy ugyanazok az emlékeik és gondolkodási stílusaik, rutinjaik lesznek. Ez utóbbi esetben mintegy digitális ikertestvérként lehet felfogni az öntudattal rendelkező, számítógépen futó, feltöltött elmét, de ahogy az egypetéjű ikrek is külön öntudattal rendelkeznek, úgy a digitális elme öntudata is külön úton jár majd a feltöltés után.

Kurzweil és Bostrom számára e kérdésben a válasz magától értetődő, hisz lévén az elme minden megnyilvánulása, minden mentális folyamat az agyi idegsejtek külségeinek eredménye, vagyis csak a materiális folyamatok emergens (az elektrokémiai folyamatok szintjéről felbukkanó, e fölé emelkedő!) Következménye, így a tudat és az öntudat is csak ennyit jelent. Ebből az alapállásból következik, hogy ha elég pontos és részletes az emuláció, akkor nemcsak agyi folyamatok részletei (emlékek, élmények stb.) jelennek meg a számítógépes futtatás folyamán, hanem az ezek összegződéséként létező öntudat is. Abban azonban, hogy ez a gépi-elme öntudat mennyiben jelenti az eredeti megkettőződését, vagy ezzel szemben egy új létrejöttét,

nem jelenik meg náluk elemzés, és maga a kérdés is inkább csak a legújabb tanulmányokban vált középpontivá.

Több vita és a szembenálló érvek végigolvasása után én inkább hajlok annak elfogadására, hogyha elég pontos és részletes az agyi emuláció, és az egyes neuronok kapcsolódásainak trillióit is át tudják másolni a komputeres platformra, akkor valószínű, hogy a mentális folyamatok központjaként, vezérlőjeként működő öntudat is megjelenik az elme-feltöltés után. Ugyanis, ha az ember nem fogadja el, hogy az idegi folyamatok finom mintázatain kívül lenne egy külön lélek önálló szubsztanciaként, akkor csak technikai hiányosságok miatt következtethet az öntudat megjelenésének elmaradására. Ha pedig ez utóbbit kizárjuk, és a teljes elme pontos és részletes feltöltését az eddigi technikai fejlemények után lehetők látjuk, akkor nem tehetünk mást, mint elismerjük a feltöltött elme öntudatának megjelenését, mint lehetőséget. Ez azonban pusztán egy digitális ikertestvéri tudatot jelenthet az eredeti számára, de semmiképpen nem azt, hogy most már “két helyen” is megjelenik ugyanaz az öntudat, és “itt is vagyok ott is vagyok” állapot jöhetne létre. És főként nem azt, hogy az ember - megunva biológiai kötöttségét - az agyfeltöltéssel átvándorolhatna a digitális létezésbe.

Ez az álláspontja David Chalmersnek is egy frissen megjelent tanulmányában, míg vele vitatkozva Massimo Pigliucci a biológiai testhez kötött öntudat kizárólagosságát vallja. Chalmers - magát funkcionalistának, Pigliuccit biológistának nevezve - így érvel a két felfogás különbségéről: “Here philosophers divide into multiple camps. Biological theorists of consciousness hold that consciousness is essentially biological and that no biological system can be conscious. Functionalist theorists of consciousness hold that what matters to consciousness is not biological makeup but causal structure and causal role, so that a nonbiological system can be conscious as long as it is organized correctly.” (Chalmers 2014:104). Ennek a jelenleg még csak filozófiai jellegű vitának azért is van jelentősége, mert ma még - és az ezzel intenzíven foglalkozó kutatók szerint még jó pár évig - az agyi emulációnak csak destruktív technikai létezik, melyek az állatkísérletek felhasználása révén (most eltekintve egyes állatvédő csoportoktól) nem jelentenek gondot az agyi emulációk létrehozásában és a fejlesztésében. Ám mivel már jelenleg is felmerült a vitákban, hogy végső stádiumban lévő, gyógyíthatatlan betegek számára ezzel tegyék lehetővé majd az esetleges fennmaradást - akik számára az agyfeltöltés destruktív jellege már nem jelent problémát -, így fontos kiemelni, hogy ezen az úton is csak egy digitális ikertestvér teremthető meg, de az eredeti személy elenyészése így sem kerülhető el.

Egy fontos nehézséget jelent az emberi agyi emuláció megvalósításában a legújabb kutatások szerint, hogy a sok reménykedéssel szemben nem elégséges pusztán a magasabb mentális folyamatok elkülönítése és ezek emulálása, félretolva a pusztán mozgási idegfolyamatok és más testrészekkel összekötött agyi folyamatok moduláris részeit. (Például a neocortex feltöltésére koncentrálva a kisagy és a cerebellum ideghálózatának félretolásával, sokkal könnyebben meg lehetne valósítani az emulációt.) E reményekkel szemben ugyanis az agykutatások azt mutatják, hogy az agy egyes régióinak specializálódott szerepjátszása mellett a legtöbb idegi folyamatba minden agyi régió belejátszik több-kevesebb fokban. Így a megfelelően részletes agyi emuláció nem állhat meg a magas szintű mentális folyamatok régióinak feltöltésénél, hanem a teljes agy, sőt azon túl még a főbb testrészek idegi kapcsolódásait is reprodukálni kell ehhez. Ennek egy további következménye így, hogy a sikeresen emulált agy működéséhez aztán szükséges egy szintén szimulált testet is kötni, mert már technikailag sem tudna működni az eredetileg ilyen testtel összefonódott agy a komputeres feltöltés után sem: “To summarize, it is misleading to refer to areas of

brain as if they were modular. Instead, it is impossible to draw strict boundaries separating, for example, the more primitive sensorimotor functions from higher cognition such as planning, judgment, decision-making, and direction of attention. Furthermore, these areas are highly integrated functionally, and anatomically, in a complex dense network. Thus, we conclude that the aforementioned “quintessentially human” (which, employing the parlance of the first section, would fall under the symbolic as opposed to the physical) functions cannot exist independently of brain structures that are devoted to world interaction and body control. Even for an uploaded individual, a body (human-like or otherwise) would remain a necessity.” (Linssen/Lemmens 2016:5).

Ki kell térni még arra, hogy amennyiben a sikeres emberi agyi emuláció valóban létrehoz a felöltött elme addigi emberi tapasztalatainak, mentális folyamatainak komputeres futtatása közben egy öntudatot (ahogy valószínűként feltételeztük), akkor a gépi intelligencia jövője is új perspektívában lesz szemlélhető. Ugyanis az, ami a társadalmi valóságról elméleti szinten nem konstruálható meg megfelelő szinten komputeres programként, az az emberi társadalomban élő és ott az ösztönök nyomása alatt, közösségi szolidaritásban létező ember elméjében gyakorlati tudásként benne van, ráadásul a magatartásirányító szintre konkretizált receptek tömegeként is. Így ennek emulálása az elmében levő tartalomként átviheti a digitális elmébe ezt is, és ha helyes a Chalmers-szel osztott feltevésünk, akkor a digitális öntudat is létrejön az ilyen társadalmi intelligenciával is rendelkező digitális elmében. Ez pedig, noha az eddigiekben már évtizedek óta külön fejlődő gépi értelemről eltérő fajtát jelent, de egyben ez is a digitális mesterséges intelligencia egyik formáját jelenti majd. Ezzel létrejön rögtön az emberi szintet elérő mesterséges intelligencia fokozata, ezután pedig - elszakadva a biológia korlátoktól és a folyamatok itteni lassúságától - az ezerszer és milliószor gyorsabb komputeres folyamatok gyorsaságát veheti fel, most már átdolgozva az eredetileg feltöltött emberi elme sok-sok részét. Vagyis az erős MI elérésének egyik új útját jelentheti az emulációval létrejövő digitális elmék továbbfejlesztésének iránya. Ez pedig eltünteti a mesterséges intelligencián belül a társadalmi valóság naiv kisgyerek szintű átlátásának a problémáját, és így a technológiai intelligenciájának féloldalasságát. Nick Bostrom ezt a lehetőséget az előbb ismertetett Szuperintelligencia c. könyvében már számba vette, és ő is jótékony hatásuként jelzi, hogy ezen az úton az emberi elmével együtt a fejlett morális érzék tapasztalatai és társadalmi átlátása is bekerül a gépi értelembe. Ám helyesen jelzi azt is, hogy a genetikai algoritmusokon és rekurzív önfejlesztésen, illetve önmegváltoztatási képességen alapuló erős MI akármikor eltávolíthatja működési programjából ezt a morális tapasztalatot és emberi társadalmat védő működési elveit (lásd Bostrom 2015:328-330). Így végül is az erős MI létrejöttének ez az útja sem jelenthet végső biztonságot.

II.3. A genetikailag feljavított szuperintelligencia kérdése

A human géntechnológia általi intelligencia-feljavítás és élettartam-növelés terén három irány tűnik számomra érdemesnek a megvalósíthatóság szempontú elemzésre. Ennek ellenoldala, amit minden eszközzel én is tiltandónak gondolok az emberklónozás, illetve a DNS géntechnológiai módosítása révén ember és állat vegyítésén alapuló kimérák létrehozására irányuló kísérletek. E három felsorolással: 1) az embrióselekción az intelligencia növelése céljából, 2) a nanobotok véráramban végbemenő működésének megteremtése, illetve biztonságossá

tétele a belső szervek megújítása céljából, 3) és végül az agyi interfészek révén az ember biológia alapú értelmének felerősítése. Jelezni kell, hogy a mesterséges intelligencia filozófiai kérdéseivel foglalkozó szűkebb elméleti körein túl az általánosabb elemzésekkel foglalkozó filozófus és társadalomtudós körökben ezek az aspektusok már figyelmet kaptak, és az első viták már lefolytak, szemben az előbb tárgyalt, erős MI témájának és az emberi agy számítógépes feltöltésének kérdéseivel. Peter Sloterdijk vetette fel 1999-ben - a humán géntechnológia addig elért fejlődését alapul véve -, hogy az embernemesítés ezen az úton is fokozható a jövőben, és erre Jürgen Habermas reagált a felháborodás és a morális elítélés érveinek és jelzőinek legnagyobb intenzitású igénybevételével. A németeknél e kérdés érdemi vizsgálat nélküli és pusztán morális felháborodás szintjén való vitatása persze sajátos történelmi örökségükből fakad, miután az emberi eugenika korábbi törekvéseit a náci hatalmi rendszer karolta fel, és ezzel a német szellemi elit számára a legmélyebben diszkreditálta és tabuvá tette ezt az egész kérdéskört (A vita elemzéséhez magyar nyelven lásd Kiss Lajos András cikkét 2003-ból). E vitán túl az elutasítás hangvételében ugyan, de a normatív beszűkültségtől mentesebben Francis Fukuyama is foglalkozott a humán biotechnológia lehetséges következményeivel 2002-es könyvében, mely alig egy év múlva magyar fordításban is megjelent. Az elutasítás a válasz erre nála is, és ennek fő oka nála az, hogy az eddig is meglévő egyenlőtlenség az emberek és társadalmi csoportjaik között ennek révén tovább fokozódik. Ezzel a "géngazdagok" csoportjai a jövőben így már nemcsak a nagy vagyonokat és kedvezőbb életfeltételeket hagyják tovább utódaikra - szemben a szegényebb rétegek gyerekeinek siralmas jövőképevel -, hanem genetikailag is feljavított formát biztosítva számukra a társadalomban létrehozzák a "géngazdagok" és a "génszegények" társadalmi csoportjait és a minden eddiginél nagyobb társadalmi egyenlőtlenséget. (Fukuyama 2003:208-210).

Fukuyama egyenlőtlenség-növekvés félelmével szemben számomra meggyőzőbb és megalapozottabb Kurzweil állítása, mely abból indul ki, hogy a kezdeti nagyköltésű humán biotechnológiai eljárások rövid időn belül olcsóvá válnak - hisz alig van bennük anyag- és energiaköltség -, így egy idő után a legszélesebb körben bevetté és rutin eljárásokká válhatnak ezek. A kezdeti egyenlőtlenség növelés után így ez a hatás nem kell, hogy eljessze az elemzőt a támogatástól. Ezentúl még azt kell szem előtt tartani, hogy ha egy technológia nem árt a másoknak, csak az igénybevevőt felemeli a mások szintjéről, akkor az az elit tagjai között és velük kezdődően mindenképpen el fog terjedni, így pusztán a sopánkodás ágálni ellene. Egyenlőségpártiként inkább azt tegye az elemző, hogy sürgesse az állami felkarolást e témában a szabályozás szintjén, illetve a szegényebb rétegek számára az állami támogatások biztosítását az ilyen eljárások igénybe vételére. A Bostrom által elemzett, összejtekéből történő sejtátalakítások útján való többszörös embriószelekciók, amikor egy-két év alatt több generációnyi szelektálás és sokszoros intelligencianövelés érhető el, tömegesebben az ilyen tagokból álló társadalmi közösségek az értelem intézményesítésének a csúcsát jelenthetik majd. Számomra a probléma inkább abban áll, hogy mivel ez technológiailag már ma is lehetővé vált, ezért az a társadalom vagy társadalmak, melyek gyorsabban lépnek erre az útra, behozhatatlan előnyt szereznek az ettől elmaradókkal szemben. Egy ország szellemi elitjének inkább arra kellene hatnia, hogy tagjai körében mind ismertebbé váljanak ezek a lehetőségek, és nyilvános vitákban arra ösztökélni az állami szabályozást, hogy támogassa az intelligencianövelést célzó embriószelekciók igénybe vételét

A másik út, a nanobotok révén a véráramban végbemenő szervcserék és állandó megfiatalítások az előbbihez képest még inkább csak a kutatások és az

állatkísérletek szintjén kezdődtek el, de a Kurzweil és Bostrom által leírtak, illetve az eddigi tapasztalatok a technológiák terjedése és megvalósulása terén kevés kétséget hagynak afelől, hogy radikális áttörések lesznek itt is a következő években. Az előbbi fenntartások nélküli támogatásával szemben e téren már nagyobb szkepszissel kell élni, mert a pusztán élettartam meghosszabbítás az egész személyiség, habitus időskori megmerevedése mellett még csak növelheti a lét nehézségét. Már ma is inkább az a fő gond, hogy a 80-90 éves élettartamig meghosszabbított élet értelmetlenségével kell állandóan szembesülni, és képzeljük el ezt egy 130-150 éves korig kinyúló élettartam esetén (még ha néhány évvel ki is lehet nyújtani a személyiség és habitus rugalmasságát).

Az agyi interfészek jövőbeli intelligenciafeljavító fejlesztéseire rátérve, ezek ma kutatásokban és állatkísérletekben már léteznek, például egérkísérletekben a memóriafeljavító hippocampus-chipek beültetése működőképes volt, és az embereken alkalmazás kísérletei is megkezdődtek már, eleinte még például az Alzheimer kór gyógyításának rövidebb távú céljával. Ezek exponenciális gyorsasággal fejlődése - együtt a többszörös embrióselektiók tömegesebb elterjedésével - valóban az eddigi történelemben példátlan intelligencianövekedést hozhat létre a következő évtizedekben.

Összességében így a human biotechnológia eszközei útján való intelligenciafeljavítás alapvetően üdvözlendő lehet – az e téren jelzett két szigorú tiltáson túl -, és a mesterséges intelligencia társadalomba bevonásának elsősorban támogatandó útjának ezt kell tartani, mint ahogy a téma egyik kutatója írja, ezen irány mellett letéve a voksot a másik kettővel szemben: “Thus, the best approach to life extension and consciousness expansion might lie in our own marvelously complex and entire bodies, meshed with and augmented by tiny bionan machines that become a part of us, rather than the opposite vision of human migration into a machine substrate. You might grow own eternal, artificial self as you gradually become bionic, in stages so tiny that you do not even notice.” (Goonan 2014:198).

Függelék

*Hogy alkossunk meg egy tudatot. Az emberi gondolkodás titkának feltárása
(Ray Kurzweil: How to Create a Mind. The Secret of Human Thought Revealed)*

Az anyag evolúciója:

- Az atomok, különösen a szénatomok, melyeknek gazdag információs struktúrájuk négyirányú kötődésekkel egyre komplexebb molekulákká fonódtak össze. Ezzel a fizika felett létrejött a kémia világa.
- Egymilliárd év után egy különösen komplex molekula fejlődött ki, melyet DNS (angol szóhasználatban: DNA) nevezünk, és ez pontosan kódolja az információk hosszú sorát, melyek mentén egy organizmus kifejlődik, mintegy a DNS-programja valóságban való kifejtésével. Ezzel létrejön a kémia felett a biológia világa, az élő organizmusok világa.
- Egy felgyorsuló továbbfejlődésben az organizmusok kifejlesztnek egy kommunikációs döntési sejt-hálózatot önmagukon belül, melyet idegrendszernek nevezünk. Ez már tudja koordinálni az egyre komplexebb testű organizmusok egyes részeit, és ezek viselkedését a túlélés érdekében.

- Egy fejlődési szint után az idegrendszer sejtjei, a neuronok egy agyat hoznak létre, mely már intelligens viselkedést is lehetővé tesz. Ezzel létrejön a biológia felett a neurológia világa.
- Következő lépcsőfok az emlősök agyának kifejlődése, mely a többi állathoz képest egy további adalékkal rendelkezik, melynek lényege, hogy a különböző elemeket egy mintába szervezi (pattern), mely mintát egy szimbólummal azonosítja, és ettől kezdve ezt a szimbólumot használja, mint egy elemet még komplexebb konfigurációk egysége hozására. Ez a képesség az agy részei közül a neocortex révén jön létre, és ez a képesség az emlősök között az ember szintjén ér el egy olyan szintet, hogy ezeket a mintákat már olyan szimbólumokkal is rögzíteni tudjuk, mint az ideák, az eszmék, a fogalmak. Az eszme-szimbólumok rekurzív összeszövésével, átfogóbb egészbe illesztésével az ember tudássá formálja ezeket a minta-aggregációkat. Csak az ember az, mely már tudás-elemeket képes egymással összeszőni, és egyre komplexebb valóság-reprezentációkat tud velük kezelni.
- Az ember ezzel az emelt szintű agyképességével és a kezén levő visszahajlítható hüvelykujj kialakulásával már lényegesen módosítani tudta a körülötte levő fizikai világot. Ez szerszám-készítést tett lehetővé, és ezzel a neurológia szintje kibővült a technológia szintjével.
- Ezen a szinten a beszélt nyelv kialakulásával lehetővé vált a szimbólum-reprezentációk tagolt elnevezése és kiejtése, illetve egy közösségen belüli tanulásként továbbadása.
- Ezt követően a beszélt nyelv írott nyelvvé való kiegészülése az egyes szimbólumok és eszmék rögzítését és továbbadását könnyűvé tette. Ezzel már az írott nyelv könyvtárai egészítették ki az agy képességét a tudás megtartására, és inentől ezek váltak középponttá a strukturált eszmék rekurzív önmagára irányítására, és ennek révén a további tudás kinyerésében a valóság reprezentációja céljából.
- Vita van arról, hogy pl. a csimpánzok is rendelkeznek-e a hierarchikusan szervezett szimbólumok egymás közötti kommunikációjának képességével, és egy primitív nyelven a szimbólumok kiejtésével. Úgy tűnik, hogy a legközvetlenebb fizikai környezeti reakcióikban ez létezik náluk is, de ez még annak is csak töredéke, mint ami a korai előembereknél létezett.

Ez a könyv a tudat minta-felismerő elméletét fejti ki (Pattern Recognition Theory of Mind - PRTM), mely leírja a neocortex alapvető algoritmusát. A neocortex felelős az agy részei közül az érzékelésért, a memóriáért és a kritikai gondolkodásért, és az emberi agy tömegének 80%-át teszi ki. Azonban a neocortex igen nagy mértékben redundáns módon épül fel egymást milliószor ismétlődő mechanizmusokkal. Ezért ha a sokszoros ismétlődéseket leegyszerűsítjük, akkor nem is olyan bonyolult ez: "There is more complexity in a single neuron than in the overall structure of neocortex." Így fontos, hogy az agyi redundancia miatt megkülönböztessük a valódi komplexitást a látszólagos komplexitástól, mely utóbbi az egyszerű struktúra redundáns ismétlődéseiből áll.

A DNS felfedezése: Friedrich Miescher (1844-1895) fedezte fel először a sejtállományban ezt, mint egy különösen molekulát, majd Nyikolaj Koltcsov (1872-1940) írta le először ezt 1927-ben, mint a sejtekben levő, óriási, örökséget tároló molekulát, mely két tükörszálból áll és replikálódik mindig azonosan. Végül 1953-ban James D. Watson (1928-) és Francis Crick (1916-2004) írta le először pontosan a DNS szerkezetét, mint egy dupla csavarodó hosszú szálat, egy spirált. Attól függően, hogy

milyenek a DNS-en belüli bázis-párok, válik az adott organizmus egy fűszállá vagy egy emberré.

(*A fénysebességgel mozgás révén az idő lelassulása*): Albert Einstein (1879-1955) ezt 26 évesen írta le 1905-ben. Ha valami fénysebességgel mozog, akkor az idő folyása ebben a sebességi állapotban nála leáll. Einstein egy másik fontos meglátása az atombombát tette lehetővé. Azt vette alapul, hogy a skót fizikus, James Clerk Maxwell 1850-ben kimutatta, hogy a fény részecskéi - a fotonok - semmilyen tömeggel nem rendelkeznek, de mégis van lendületük, mozgóerejük. Ez indította be Einstein képzeletét, hisz a lendület, a mozgóerő nem más mit a tömeg szorozva a sebességgel. (Így egy kistömegű anyag nagy sebességgel ugyanolyan mozgóerőt ad át, mint egy nagytömegű, de kis sebességű anyag.) Ám akkor hogyan lehet mozgóerő-átadás egy tömegnélküli fotonnál? - tette fel a kérdést. Einstein felfedezése ebből az volt, hogy mivel a tömegnélküli foton energiát tud kifejteni egyszerűen óriási sebessége révén, és így ad át lendületet, mozgáserőt, akkor a fénysebességű mozgás energiája is megfelel a tömeg energiájának. "The energy of the moving photon is entirely equivalent to a moving mass." Mi ez az ekvivalens – kutatta -, és matematikailag számolta ki: ha egészen kis részecskék fénysebességű mozgással robbannak fel, akkor egy 28 grammos atommaghasadás részecskéi a fénysebességű mozgással kb. 600 ezer tonna TNT erejének felelnek meg - számította ki. Vagyis, ha egy kis anyag-tömeg részecskéit fénysebességre tudjuk villanásnyi idő alatt felgyorsítani, akkor egy atombombát kapunk - ez volt a belátása. Ekkor Einstein írt egy levelet Roosevelt elnöknek 1939 aug. 2-án, és felhívta a figyelmét egy atombomba kifejlesztésének lehetőségére. {Máshol úgy írták, hogy erre más tudósok – köztük Neumann János - beszélték rá Einsteint, és ő csak ehhez csatlakozott. }

Chapter 3.A neocortex (magyar szóhasználatban: új agykéreg)

- A neocortex nélküli állatok (a nem emlősök) képtelenek megérteni a hierarchikus elrendezést.
- A neocortex felel az érzéki észlelésért, a vizuális észleléstől kezdve az absztrakt eszmék létrehozásáig.
- Az emberi neocortex az agy legkülső rétege, és egy vékony, két dimenziós struktúrában épül fel, kb. két és fél centis vastagságban. A rágcsálóknál a neocortex csak postai bélyeg nagyságú, de a főemlősöknél már felveszi lassanként azt az alakot, ahogy azt az embernél láthatjuk. A emberi agy 80%-ának a súlyát a neocortex adja. A homo sapiens fejlesztett ki olyan homlokrészt a koponyájában, amely lehetővé tette a még nagyobb neocortex növekedését a fejben. Különösen a neocortexnek a homok alatti részében történik az absztraktabb mintákkal való foglalkozás.
- Ez a vékony neocortex struktúra maga is hat vékonyabb rétegből áll. A legkülső a VI. réteg. Az outputok a II. és a III. rétegből indulnak, melyek a neocortex más részei felé irányulnak, és ezek az output-axonok az ottani input-dentritekhez kapcsolódnak. Ezzel szemben a V. és a VI. rétegek axonjai a neocortexen kívüli agyrészek felé kapcsolódnak, különösen a talamus felé, a brainstem (agytörzs) illetve a spinal cord (gerincvelő) felé. Végül pedig a IV. rétegben levő neuronok a neocortexen kívüli agyrészekről inputokat kapnak. Vagyis míg belülről kifelé irányulóan az V. és a VI. neocortex rétegek csatlakoznak, addig kívülről befelé a IV. réteg neuronjai fogadják a csatlakozásokat. Külön kell a IV. rétegben említeni a neocortex occipital lobe részét, mely a vizuális folyamatokért felelős, és három álréteggel rendelkezik, és a talamustól ide jönnek ezek az inputok.

- A neocortex rendkívüli mértékű egységes felépítését már 1957-ben felfedezte Vernon Mountcastle, és ezt 1978-ban pontosan ki is fejtette. Leírta, hogy a neocortex felépítése egyetlen mechanizmusból áll, amit újra meg újra megismétel a neocortex egyes részein. Ez egy elementáris alapegységet jelent, egy agyi pillért, mely állandó ismétlődéssel sokszorozódik a neocortex összes részén. Minden elementáris pillérben ugyanazok az ismétlődő egységek vannak, és Kurzweil állítása, hogy ezek a basic unit-ok a neocortex mintafelismerő egységei. A pilléreken belül ezek egymás mellett sorakoznak, mint idegsejtek. Ezek a minta-felismerő idegsejtek pedig össze tudnak kapcsolódni egymással egész életen át, és ez a kapcsolódás nem eleve létezik a genetikai kód által meghatározva, hanem tanulással-tapasztalással jönnek létre, és ezen összekapcsolódások révén tárolódnak a tanult információk. Ezek a tanultak, a tapasztalások a neuronok fiziológiai szintjén aztán nem mások, mint a neuronok összekapcsolódásának egy-egy mintázata.
- Az emberi neocortexben mintegy félmillió ilyen agyi pillér (column) van, és mindegyik kb. 2 milliméter magas és fél milliméter vastag, és egy-egy ilyen pillér kb. 60 ezer neuront tartalmaz. Összesen így kb. 30 milliárd neuron van az emberi neocortexben, ám mivel kb. száz neuronból áll egy-egy mintafelismerő mechanizmus, így kb. 300 millió az összes mintafelismerő száma az emberi neocortexben. Ezt a 300 millió mintafelismerőt teszi a középpontba Kurzweil, amikor az emberi tudat működését próbálja leírni.
- Fontos látni, hogy a neocortexben levő 300 millió mintafelismerő idegsejt-rendszer párhuzamosan működik. Ahhoz, hogy lássuk, hány mintát tud tárolni a neocortex ezzel a 300 millió mintafelismerőjével, fontos kiemelni, hogy óriási ismétlésekkel tárolja a mintákat ezekben a mintafelismerőkben. A legkedvesebbeket, a számunkra legfontosabbakat akár ezerszer is megismételve. Ezek a sokszoros ismétlések persze különböznek a részletekben, pl. az arcot különböző megvilágításokban és eltérő irányokból is tárolja külön-külön mintaként az agyunk, ugyanígy nemcsak a egész arcot, hanem annak legjellegzetesebb vonásait is külön mintaként. Például ezért van, hogy De Gaulle nagy orra egy karikatúrán elég ahhoz, hogy rögtön tüzeljen egy mintafelismerő neuron-csoportunk, és beugrik az ő képe. Vagyis egy-egy mintafelismerő a 300 millió közül egy-egy mintára fonódik rá, az lesz a tartalma, és ha ez a minta feltűnik, akkor az a mintafelismerő tüzel, hogy felismerte, és az jelen lesz az agyunkban.
- Az agyban megjelenő minták és az ezeket felismerő mintafelismerő neuron-csoportok hierarchikusan szervezettek. Minden minta három részből áll:
 - az egyik részt az input-rész jelenti, amely az alsóbb szinten lévő mintákból áll;
 - a minta második része a minta nevét tartalmazza;
 - a harmadik része pedig azokból a magasabb szinten lévő minták készletéből áll, pl. ha az "A" betű a tartalma egy mintafelismerőnek, akkor egy felsőbb szinten a szó-minták mindegyike felé küld egy axont (outputot), melyekben az "A" betű szerepel.
- Egy-egy életbeli dolgot a mintafelismerők sokasága jeleníti meg a neocortexben. Pl. az "alma" minden formájára egy-egy mintafelismerő szakosodik, írott alakjára, vizuális alakjára, hangtani beszéd-alakjára. De még ezen túl is sok-sok mintafelismerő redundáns módon az almát jeleníti meg eltérő világításban, alakokban. Ám így lesz biztos, hogy az agyban a sokszorosán tárolt dolog felismerése bármelyik előfordulás esetén létrejön.

- Az új mintákat a neocortexen belül a hippocampus tárolja. Mint új minták - miután a felmerüléskor egyik szakosodott mintafelismerő sem ismerte fel, és nem tüzelt – a hippocampusban tárolódnak. Ez tartalmazza magát az új mintát, és azt, hogy milyen környezetben merült fel, pl. hogy neveltünk–e a látásán, amikor észleltük stb. De ha egy már ismert minta részben új kiegészítésben merül fel, akkor ez is tárolva lesz mint egy új minta, kapcsolódva részben a már eddig ismert alakjának mintájához.
- Ha a tapasztalás során egy már ismert mintát találunk, akkor a mintafelismerő neuroncsoport tüzelni fog, és a felismert minta megjelenik a tudatunkban, de az nem marad meg a hippocampus memóriában, pusztán rutin-szerű eltűnő észlelés lesz. Aminek nincs valami új sajátossága, az eltűnik, és nem válik a memória részévé új mintaként.

A tanulás idegsejt-szintű feltárása

- Donald O. Hebb (1904-1985) állította először 1949-ben, hogy a tanulást észlelni lehet idegsejti változások szintjén is. A tanulás hatására egy-egy idegsejt fiziológiailag is módosul, és ez adja a tanulás és a tapasztalás révén az agy plaszticitását. A tanulás hatására egy-egy idegsejt összekapcsolódik egy sor másikkal, és egy idő után, ha az összekapcsolódó idegsejt tüzel, akkor vele együtt tüzelnek a vele tartósan összekapcsolódott idegsejtek is. Vagyis összefűződnek egy tapasztalat tárolására egy idegsejt-együttessé. Ez a “Hebbian-tanulás” tézise.
- Kurzweil jelzi, hogy az ő mintafelismerő elmélete Hebb elméletének egy módosítása. Ugyanis nála nem egy-egy idegsejt a tanulás alapegysége, hanem eleve egy idegsejt-együttes, ami kb. száz idegsejtből van összefűződve. Ezek már genetikailag összefűzöttek, és nem a tanulás fűzi össze őket. A tanulás már csak ezeket az idegsejt-együtteseket fűzi egymással össze. Ezt megerősíti Henry Makram (1962-) svéd idegtudós nézete is, aki a *Blue Brain Project* révén az agy teljes emulációját igyekszik elvégezni, és az emberi agy teljes emulációját 2023-ra ígéri.

A vizuális agyi rész működése

- A vizuális cortex három részből áll, a V1, a V2 és az MT, vagy V5-ként is ismert.
- A V1 csak az alapvető kontúrokat ismeri fel, a sarkokat és az átfogó alakokat. A V2a teljes kontúr, a helyi orientációkat, és hogy egy rész a kép része-e vagy inkább már a háttér része. A neocortex magasabb részei ismerik fel a kontextuális részeket, a teljes alakok identitását és a konceptuális részeket illetve ezek mozgását.
- E felismerés közben állandó kommunikáció folyik a részek között, és pl. egyes részletek felmerülése, mint tiltó jelzés megy a konceptuális azonosítás felé (pl. ha egy bajuszt látott az ember, az egy feleségként azonosítást csökkentve egy tiltó jelzésként megy fel az identitás meghatározása felé.) De ugyanígy fordítva, egy részlet erősítheti egy jelzéssel felküldve az észlelést végző agyi résztől a konceptuális rész felé egy meghatározott identitás meghatározását. Tehát állandó fel- és lefelé irányuló kommunikáció folyik a vizuális agyi részek között “látás” közben, és nem csak “objektíven” látunk a szemünkkel valamit egyszerűen, hanem az egy sokszoros összetett agyi meghatározás alapján adja meg a látottakat. Tomaso Poggio (1947-) sokat tette ennek feltárásáért.
- Az agy “fusiform gyrus” része felelős az arcfelismerésért, és ha ez károsodik, akkor az ilyen ember nem tudja arcról felismerni az egyéneket, de hangjukról és

beszédstílusukról még igen. Ám egy idő után az agy egy másik része kialakít egy új központot az elromlott képesség pótlására. Ez az agy plaszticitása, és egy másik rész átveheti bizonyos fokig a sérült agyi rész funkcióját.

- Megfigyelték, hogy a született vakoknál a vizuális központ átveheti a beszédközpont feladatait is, és ebbe beszállva a beszédfunkciót segíti. Igaz, hogy ez is az agy bal féltekéjében van, közel ahhoz, ahol az eredeti beszédközpont van.
- A tanulás fiziológiai bázisáról ismét: Yi Zuo idegtudós és csoportja figyelte meg egy kísérletben, hogy ha az egér egy új tudást tanult, akkor ez bizonyos idegsejtjei között stabil kapcsolódást hozott létre. “A new ‘dendritic spines’ formed connections between nerve cells as mice learned a new skill (reaching through a slot to grab a seed) (...) an actual physical connection must be made composed of an axon connecting to a dendrite.”

Chapter 5. - *The Old Brain (a magyar szóhasználatban ez a kisagy)*

- Az öreg agyat a mai ember még akkor alakította ki, amikor még nem érte el az emlős szintet, és ezt azóta is megtartotta. Ez felelős még ma is egyes motivációink, a félelem- és gyönyörérzésünk alapjaként. A nem emlős állatoknál még ma is csak ezzel élnek, neocortex nélkül. Az emberi neocortex együttélését az öreg aggyal úgy kell elképzelni, mint a nagy megnemesítőt. Az ember ugyanis ezzel szublimálja rögtön az öreg agy eredeti reakcióit, és elrejtje ezeket az eredetileg primitív reakciókra ösztönzéseket a kifejtésük előtt. Ezek a reakciók arra voltak szabva, hogy ha a csak ilyen aggyal rendelkező állat meglát egy számára ragadozót, akkor ösztönösen beugrik a reakció, hogy meneküljön.
- (*Az érzékelés útvonalai*) Mi ugyan úgy véljük, hogy a szem által látott képeket egyből a külvilágból kapjuk, de valójában a látási idegsejtek küldenek előtte az agynak sok-sok vázlatot, vonalat, alakot, kontúrokat a vizuális mezőnkbe kerülő dologról, és az agy állítja össze ezekből az agyban levő minta-felismerőkkel (ténylegesen tehát a memóriánk hozzáadásával) a képet. Egy kísérlet kimutatta, hogy 10-12 idegi csatornán keresztül hozzák létre az idegsejtek a végső összkép egy-egy részletét. Pl. a ganglionsejtek a sarkokról és a kontrasztokról hozzák az információkat, egy másik látóideg csoport az egységes színek területeit azonosítja, egy harmadik pedig csak a háttérrel látja az alak mögött. És ezekből állítja össze végül az agy vizuális központja az egész képet.

A talamus

- Az érzékelés elsődleges inputjai után a hallás, a látás, stb. információi a középagytól átmennek a talamus egyik részébe, a *posterior ventromedial nucleus*-ba (VMpo), és ez a testi észleléseket egységes reakciókkal látja el. Például “ez az íz szörnyű!” vagy “micsoda büz!” Ez az összeadó információ éri el aztán a neocortex két régióját, melyet *insula*-nak nevezünk. Ennek a struktúrái – mintegy a kisujj méretének felel meg - a neocortex jobb- és baloldalán vannak. Arthur Craig úgy nevezte a neocortex kettő *insula* részét és a VMpo-t mint az Én-ek a materiális részét.
- A tapintás-érzékelés hasonló utakon megy át az érzékelőtől a gerincvelőn át az agytörzsbe (brainstem) majd a VMpo-n és az *insula*-n át a neocortexbe. De itt a *lateral geniculate nucleus* rész az, ami bejön a feldolgozásba.

- A hallásnál a thalamus *medial geniculate nucleus* része jön be feldolgozóként: “All of our sensory data (...) pass through special region of the thalamus.” De a thalamus legfőbb feladata, hogy minden érzékelés-továbbítás közben kommunikál a neocortex-szel. A neocortex mintafelismerő jelzést küld a megkapott érzékelésre válaszként a thalamusnak, hogy milyen minta várható az érzékelés alapján. A jelzés-küldés ténylegesen idegfonalak összekapcsolódása révén jut át a neocortexből a thalamusba axonok formájában, melyek ott a dentritekre kapcsolódnak rá. Így a 300 millió mintafelismerő valamelyike a neocortexből állandóan kapcsolódik egy-egy érzékelés nyomán a thalamushoz.
- Akinek megsérül az agy thalamus része, annak a neocortex része még tud meditálni (self-triggering thinking) asszociálva a benn lévő tartalmakra a memóriában, de a külvilág felé az események iránt öntudatlan állapotba kerül. Például a 21 éves Karen Ann Quinlan egy szívszélütés után öntudatlan állapotba került, és tíz évig már csak vegetált, ám amikor meghalt a boncolás kimutatta, hogy a neocortex része egész addig működött, és csak a thalamus része volt szétroncsolódva.
- A thalamus fő feladata, hogy irányítja a figyelmet aszerint, hogy a neocortexből milyen jelzést/utasítást kap. Az ember csak agyának thalamusa és neocortexe párhuzamos működése révén tud normálisan létezni.

A hippocampus

- Mind a két agyféltekének van egy hippocampus része, amely egy kis terület az agyban, és olyan az alakja, mint egy tengeri csikónak. Fő funkciója, hogy az új eseményeket tárolja, mint emlékeket. Mivel a az érzékelések átmennek a külvilágról a neocortexbe, az ottani mintafelismerők döntenek el negatívan, hogy új, nem ismert mintáról van-e szó, és ha igen, akkor ez átkerül a hippocampusba tárolásra. De ha a mintafelismerő valamelyike fel is ismeri, ám egy új részlet van a részletekben akkor, is mint új minta bekerül tárolásra a hippocampusba.
- Az neocortex nélküli aggyal rendelkező állatok az érzékelteket egyből a hippocampusukba küldik, és ez csak érzékelésként tud emlékké válni. (A neocortex-szel rendelkezők esetében - az embernél főként - az emlékek tárolása a neocortex jeleivel együtt kerül át a hippocampusba!) Ezért a neocortex nélküli aggyal rendelkező állatok a tapasztaltakat nem tudják rendezni hierarchikus tapasztalásokká.
- A hippocampus kapacitása kicsi, és így az itteni memória rövidtávú csak, és a tárolt memóriát időről - időre átküldi a neocortex hosszútávú memória részeibe. Theodor Berger, kaliforniai idegtudós kollégáival végzett kísérleteket e tárgyban patkányokon. Be tudott ültetni az agyukba egy mesterséges hippocampust 2011-ben, miközben drogokkal blokkolták a patkányok tanult magatartását. A beépített hippocampus révén azonban a patkányok ismét meg tudták tanulni a blokkolt viselkedési mintáikat, és így újra tudtak reagálni. Ekkor, ha lekapcsolták a beültetett hippocampust, akkor az újratanulás után a patkány hirtelen ismét elfejeltette ezt, de ha visszakapcsolták, akkor ismét feléledtek az emlékei. Egy másik kísérletben a beültetett hippocampust együtt működtették a patkány meglévő hippocampusával, és ez sikeres volt, a patkány tanulási képessége (új emlékek tárolása) felerősödött. “The hippocampus is one of the first regions damaged by Alzheimer’s, so one goal of this research is to develop a neural implant for humans that will mitigate this first phase of damage from the disease.”

A cerebellum (a testi mozgás koordinációját végzi)

- A cerebellum a régi agy részét jelenti, amely valaha az összes mozgást koordinálta, de az embernél már a neocortex is bejön e feladatok meghatározásába. Az agy összes idegsejtjének fele itt található meg, de ezek nagyon kis méretűek a többihez képest, úgyhogy az agy teljes súlyának csak tíz százalékát teszik ki. Helyileg a cerebellum a brainstemre (agytörzsre) van rátelepedve a gerincvelő feletti részen, és egy baseball-labda nagyságú bab alakú rész. Az embernél már az összes izommozgást a neocortex ellenőrzi, és ugyanúgy a mintafelismerőket használja ehhez, mint az észlelésre és gondolkodásra is, de a cerebellumban lévő memóriát használja fel ennél. Ez a memória a felismert minta (dolog) mozgási jellegzetességeit tartalmazza, például bizonyos írásmozdulatok jellegzetességeit egy személy esetében, vagy ugyanígy egy művészeti alkotó jellegzetes mozdulatait, pl. festői vonásokat, vagy táncbeli mozgásokat. Tanulmányok a gyerekek kézírásának tanulásáról feltárták a cerebellum Purkinje-sejtjeinek szerepét ennél. Ezek összegyűjtik az adott írás-mozdulatok szekvenciájához tartozó mozgásokat. Mivel a mai ember már a legtöbb mozdulatát a neocortex-szel ellenőrzi-irányítja, ezért még komoly cerebellum sérülések esetén is tudja produkálni az ilyen sérült ember a mozdulatokat, de ez már nem lesz olyan kecses, mint a cerebellum irányítással együtt korábban volt.
- A cerebellumban van egy született képesség, hogy kiszámítsa intuitív módon a mozgások következményeit, amelyet már tervbe vett (a neocortexe) de még nem hajtott végre. Illetve ugyanígy, a partnerek lehetséges későbbi válaszmozdulatainak következményeit, hogy így tegye meg már előre az ő mozdulatait. Ilyen a született boxolói képesség és a szeniális focisták cselezési képessége is.

Az amygdala és a nucleus accumbens (ezek a félelem és a gyönyör-érzés helyei)

- Tehát a félelem-érzés az amygdala agyi részben alakul ki, itt tüzel a neuron, ha az ember agyában ilyen érzés alakul ki, míg a gyönyör érzése az agy nucleus accumbens részében jön létre.
- A legkorábban kifejlett agyi részek ezek, és az egyszerű állatok ezzel reagálnak a számukra fontos élelemszerzési szükségletekre, illetve veszélyhelyzetben, hogy ne váljanak prédává. Ám az emberben is megvannak még ezek a korai agyi reakciók és agyi részek, mégha teljesen ellenőrzés alá vonva is a neocortex által. De ha nyersen tör elő a félelem érzete vagy a gyönyör érzése, akkor az embernél is a régi agy amygdala része, illetve nucleus accumbens része produkálja tisztábban ezt az érzést. (Például patkány-kísérletekben a gyönyör érzésért felelős nucleus accumbens agyi részt izgatva elektródákkal vég nélkül csak evett vagy szexelt a patkány, míg el nem pusztult.) Az embernél a gyönyör érzésénél még más agyterület is közreműködik, amely involválódik a gyönyör keletkezésénél, ez a *ventral pallidum*, de maga a neocortex is részt vesz bizonyos fokig ebben.
- A gyönyör érzése kemikáliák segítségével is létrehozható. A *dopamin* nevű kemikália egy olyan idegstimuláló, amely a gyönyör érzését kelti fel. Bármilyen örömről és jó érzésünkről van (pl. nyerünk a lottón, kapunk egy jó hírt szerett társunkról stb.) dopaminkibocsátást lehet mérni ezeken az agyi részekben. Ha egy ember hozzácsókol a dopaminhoz, akkor a speciális dopaminreceptor, a D2 gén mutálódik, és ez aztán különösen erős gyönyör-érzést hoz létre, ám ezután az első érzés után mindig radikálisan csökken ez, és szinte teljesen vége a gyönyör-érzésnek. De ugyanígy káros az is, ha valakinek soha nincs dopaminkibocsátása, és így lesz genetikai torzulása a D2

génjében. Ez egy kisebb csoport a társadalmakban, és ez is szociális és egészségügyi gondokat hoz létre.

- Egy másik kemikália a *serotin*, amely pedig a hangulati állapotért felel az agyban. Magas serotin szint a jó érzés állapotát hozza létre, de a serotin szintünk szabályozza az étvágyat, az alvási készséget, a szexuális vágyat és az emésztést is. Az anti-depresszáns gyógyszerek szelektíve a serotin szintet szabályozzák, növelik a serotin szintet oly mértékben, amely még elfogadható a serotin receptorunk által. A probléma azonban ezekkel az, hogy míg a neocortex idegsejt-akciói csak a neocortex kis területét érintik, addig e kedélyállapot- illetve gyönyör fokozószer (serotin, dopamine) az agy teljes területére kihatnak, sőt az egész idegrendszerre. (Máshol beszúrva: a dopamine magas szintje (és a norepinephrine) rövid távon nagy figyelmet és eufóriát tud kiváltani, és így nyilván jelen vannak ezek a szerek a sport-dopping szerekben.)
- Az amygdala mindkét agyfélteke részen jelen van. De ugyanígy a régi agyban is benne van, annak az egyik része, és egy sor érzést ez állít elő, különösen a félelem érzését. Ha olyan szituáció áll elő, amit életveszélyként fogunk fel, akkor az amygdala létrehozza a “menekülj vagy küzdj!” régi páni félelem-érzés reakcióját. Az embernél persze ezt ellenőrzi még a neocortex is, és ez legtöbbször megszelídíti ezt az eredeti reakciót: “Emotional experiences take place in both the old and new brains. Thinking take place in the new brain (the neocortex) but feeling take place in both. (...) There is a continual struggle in the brain as to whether the old or the new brain is in charge. The old brain tries to set the agenda with its control of pleasure and fear experiences, whereas the new brain is continually trying to understand the primitive algorithm of the old brain and seeking to manipulate it to its own agenda.”
- A “*spindle neuron*”-ok is az újonnan létrejött agyi struktúrát jelentik – túl az amygdala régi jellegén -, és a magasabb rendű érzések főként itt képződnek. Szemben a neocortexben található szabályos és rekurzív jellegű struktúrákkal a spindle neuronok szabálytalan alakúak és kapcsolódásaik is ilyenek. Ezek az emberi agy leghosszabb idegsejtjei, és az agy egész hosszát végigérik. Nagyon össze vannak egymással fonódva, és a kapcsolódások százezreivel vannak összefonódva, a neocortex legkülönbözőbb területeibe beelölőva.
- Az insula, amely az érzékelési jeleket továbbítja, szintén részt vesz a magasabb rendű érzések működésében. A spindle sejtek itt az insulából erednek, és mágneses kísérletekkel meg lehetett állapítani, hogy ezek a spindle sejtek nagyon aktívvá válnak, ha egy magasabb rendű érzés létrejön, pl. szerelem, szomorúság, vagy szexuális vágy, harag, düh. A spindle sejteknek hosszú idegszálaik nyúlnak ki, ezek az *opical dendrite*-ek, melyek távoli neocortex részeket is összekapcsolnak. Az evolúció magasabb fokán fejlődtek ki ezek a több neocortex részt összefogó idegszálak. Ám épp e mindenhol való jelenlétük miatt a spindle sejtek minden érzékelésnél és kognitív akciónál jelen vannak, és érintve vannak ezek által. Ezek az idegsejtek nem vesznek részt a racionális problémamegoldásokban, így ha pl. szerelembe esünk, vagy a zene elérzékenyít, akkor nem tudunk racionális lenni ez alatt.
- Kb. 80 ezer ilyen spindle sejt van az emberi neocortexben, 45 ezer a jobbféltekében és 35 ezer a balban. Ez a különbség okozza, hogy az emocionális intelligencia a jobbféltekében van. A gorilláknak 16 ezer ilyen sejtje van, a bonókó majmoknak 2.100, a csimpánzoknak 1.800. Más emlősöknél ez teljesen hiányzik. Kb. 10-15 millió évvel ezelőtt jelentek meg ezek a spindle sejtek a majmok és az emberféléknél, közös őseinknél, és kb. 100 ezer évvel ezelőtt radikálisan elkezdett nőni a számuk az akkori emberek agyában. Az újszülöttek agyában ez nincs még meg, és csak 4 hónapos korban kezdenek megjelenni, majd 3 éves kortól kezdenek erősen sokasodni.

Alkalmassági képesség, született rátermettség

- Mozart öt évesen már írt egy menüet, és hat éves korban zongorázott Mária Teréziának Bécsben. Ez a született képesség kiemelkedő példája. A nevelés is számított de a születése már megadta genetikailag az alapot ehhez. Apja, Leopold Mozart karmester és zeneművész volt Salzburgban a püspöki zenekarban, és már három évesen elkezdte a fiát tanítani hegedülni és zongorázni.

VII. fejezet - A biológiailag inspirált digitális neocortex

Az agyi szimuláció

- David Dalrymple (1991-) Harvard agytudományi doktori hallgatója egy féreg agyát tervezi szimulálni (roundworm), mely csak 300 neuronból áll, és ezeket a molekulák szintjén igyekszik szimulálni. A kutatás másik végén található Henry Markram Blue Brain Projectje, aki a teljes emberi agy szimulálására készül, a neocortex-szel és az összes régi agyi résszel, benne a hippocampus, az amygdala, a cerebellum. Ő több szinten kívánja ezt a szimulációt végrehajtani, bizonyos részeken egészen a molekuláris szint részletességével, de máshol csak átfogóbb funkcionális szinten. Felfedezte, hogy több tucat idegsejt összefonódva egy modulát alkot, és a tanulás ezeken a modulok szintjén megy végbe. Vagyis az agy és nem az egyes idegsejtek szintjén.
- Markram és csoportja 2005-ben egy neuront tudott még csak szimulálni (ez az év volt a kutatása kezdő éve), és 2008-ra már egy patkány neocortexének egy egész modulját szimulálta, amely 10 ezer neuronból állt. Aztán 2011-ben már száz modulra terjesztette ki a szimulációt, és ezek összesen már egy millió idegsejtet tartalmaztak. Ez kb. egy mesoterületnek fogható fel a patkány agyában, és a teljes patkány agy 100 ilyen mesoterülettel rendelkezik. Ez a 100 millió neuron a patkány agyában kb. egy milliárd szinapszist alkot, és Markram ezt 2014-re tervezte befejezni. Találkozott vele Kurzweil 2009-ben egy konferencián, és akkor mondta neki, hogy a teljes emberi agy emulációját 2023-ra tudja befejezni. Az optimizmusát az adta, hogy a csoportja alkotott egy nanoméretűben szkennelő robotot, és ez már addig is 30 év munkáját tudta elvégezni hat hónap alatt. Tehát ez a *patch-clamping robot* az idegszöveteket szkenneli be Makramék számára.
- Ám az a probléma, hogy a teljesen kész szimulált agy még üres lesz elkészülte után is 2023-ban, mint egy újszülött agya, mely csak kapacitással rendelkezik. Így még egy külön feladat lesz a digitális agy betanítása, ahogy az újszülöttnél is van. Vagy a fokozatos tanítással lehet feltölteni a digitális agyat, vagy egy már felnőtt agy tartalmának letöltésével, és ennek a digitális agyba való áttöltésével. Ám a non-invasive és non-destructive agyletöltés még csak a kezdetek elejénél tart (a patkány agyi kísérleteknél ezzel nem kellett törődni...), így Kurzweil 2040 körülre teszi ennek a teljesen feltöltött digitális agynak a létrejöttét. De az ehhez szükséges non-destructive módszerek még ennél is hosszabb időt vehetnek majd igénybe – írja. (Közbevetve: a 2005-ös Szingularitás könyvében még fixre vette a 2040-es dátumot, mint legkésőbbi időpontot ehhez, így itt már egy módosítás van ezen!)
- Egy másik agyletöltési project Dharmendre Modha csoportjéé az IBM-nél. Ők sejtszintű szimulálást végeztek az emberi vizuális neocortex részeken, és ez 1,6

milliárd neuronai rendelkezik, melyek 9 milliárd szinapszisban fonódnak össze. Ez egy macska teljes agyának felel meg.

- Az oxfordi Anders Sandberg (1972-) és a svéd filozófus, Nick Bostrom írt közösen egy könyvet "Whole Brain Emulation: A Roadmap" címmel, és ebben összegezték az emberi agy szimulációjának lépcsőfokait, ill. más, kisebb emlősök agyának szimulációs lépéseit.
- Közbevetve: John von Neumann a szingularitás első megfogalmazója, de ezt csak közvetetten tudjuk Stam Ulam-tól. Ő 1957-ben, egy évvel Neumann halála után idézte fel egy írásában a vele való beszélgetését a korai 1950-es évekből, és ekkor beszélt neki Neumann röviden a szingularitás eljövételéről:
"The ever accelerating progress of technology and changes in the mode of human live give the appearance of approaching some essential singularity in the history of the race beyond which human affairs as we know them, could not continue". This is the first known use of the word "singularity" in the context of human technological history."³

Irodalom

- Bostrom, Nick: Szuperintelligencia. Ad Astra Kiadó. Budapest 2015.
- Chalmers, David J.: Uploading: A Philosophical Analysis. In: Russel Blackford/Damien Broderick (eds.) Intelligence Unbound: The Future of Uploaded and Machine Minds. Wiley Blackwell. Malden - Oxford. 2014. 102-118.p.
- Fukuyama, Francis: Poszthumán jövőnk. A biotechnológiai forradalom következményei. Európa Kiadó. Budapest. 2003.
- Goertzel, Ben/Joel Pitt: Nine Ways to Bias Open-Source Artificial General Intelligence Toward Friendliness. In: Russel Blackford/Damien Broderick (eds.) Intelligence Unbound: The Future of Uploaded and Machine Minds. Wiley Blackwell. Malden - Oxford. 2014. 90-101.p.
- Goonan, Kathleen Ann: The Future of Identity: Implications, Challenges, and Complications of Human/Machine Consciousness. In: Russel Blackford/Damien Broderick (eds.) Intelligence Unbound: The Future of Uploaded and Machine Minds. Wiley Blackwell. Malden - Oxford. 2014. 193-200.p.
- Kiss Lajos András: Az emberiség bizonytalan jövője: Habermas és Fukuyama a biotechnológus morálfilozófiai kérdéseiről. Holmi 2003 11. sz. 1469-1474.p.
- Kurzweil, Ray: A szingularitás küszöbén. Amikor az emberiség meghaladja a biológiát. Ad Astra Kiadó. Budapest 2014.
- Kurzweil, Ray: How To Create a Mind. The Secret of Human Thought Revealed. Viking Penguin Edition. London. 2012.
- Linssen, Charl/Pieter Lemmens: Embodiment in Whole-Brain Emulation and its Implications for Death Anxiety. Journal of Evolution and Technology. 2016. No. 2. 1-13.p.
- Luhmann, Niklas: Az értelem mint a szociológia alapkategóriája In: Luhmann, N.:

³ Felvetés: tulajdonképpen hosszan ezen meditálva fejlesztették ki Kurzweilék a szingularitás eszméjét fokozatosan az elmúlt években mint az emberi értelem fölé nőő és attól függetlenedett, magasabb rendű intelligencia létrejöttét, és a szingularitás átlépési küszöbét úgy határozta meg, mint amikor a gépi értelem addig együttműködő emberi értelem és ennek lassúsága végleg kiszáll a gépi értelem egyedüli működéséből, és ezzel a gépi értelem milliószeres gyorsasága megszabadul a emberi lassúságtól, és robbanás-szerű változások indulnak meg, néhány óra alatt korábbi évtizedes változásait végrehajtva.

Válogatás írásaiból. (Válogatta Pokol Béla) Szociológiai Füzetek 42.
Budapest 1986. 32-47.p.

Pigliucci, Massimo: Mind Uploading: A Philosophical Counter-Analysis. In:
Russel Blackford/Damien Broderick (eds.) Intelligence Unbound: The Future
of Uploaded and Machine Minds. Wiley Blackwell. Malden - Oxford. 2014.
119-130.p.