Pokol Béla

Elméletek az MI társadalomra gyakorolt átfogó hatásairól Kelet-Európában 2000-2026

(a ChatGPT5.6 extra erőS verziójának elemzésével)

Budapest
2026

Tartalom

Átfogó bevezető.....3

I. Magyarország.....22

Pokol Béla.....22

Csepeli György……..... 43

II. LENGYELORSZÁG.....64

Andrzej Zybertowicz.....64

Jacek Dukaj.....101

III. SZLOVÁKIA.....126

Jozef Kelemen.....126

IV. HORVÁTORSZÁG.....149

Ivana Greguric.....149

Nenad Vertovšek.....171

V. SZERBIA.....192

Vladan Joler.....192

VI. ROMÁNIA.....221

Mihai Nadin.....221

VII. Európai elméletalkotók originális tézisei és állításai az AI társadalomra tett hatásairól.....242

Átfogó bevezetés

Kérjük, részletesen írja le és elemezze magyar, lengyel, cseh, szlovák, horvát, szerb, szlovén és román szerzők évezredfordulótól kezdve végzett tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezték, beleértve az AI további radikális fejlődését is, amelyet a következő években előrejelzett. Csak azok a szerzők vehetnek részre, akik műveikben átfogó elméleti rendszert dolgoztak ki ezen a területen, és előnyben részesítik azokat, akik eredeti téziseket is dolgoztak ki ezekben az elméleti tanulmányokban.

****

Átfogó elméletek az MI és a társadalom Közép- és Kelet-Európában 2000 óta

A választást jóval szűkebbnek tenném, mint a hagyományos "MI és társadalom" bibliográfiája. Az általad említett nyolc országban ma már több száz tanulmány található az AI etikáról, foglalkoztatásról, oktatásról, jogról, orvostudományról, közigazgatásról, katonai ügyekről, algoritmikus torzításról és digitális kultúráról. A legtöbbjük azonban **nem alkot átfogó elméletet az AI társadalmi átalakulásáról**.

Ezért a jelen felméréshez öt összesített követelményt alkalmaztam. A szerzőnek viszonylag koherens képet kell adnia arról, mi is valójában az MI az emberi intelligenciával kapcsolatban; összekapcsolnia kell a technológiai változást a társadalom több fő szférájával egy szektor helyett; azonosítania egy ok-okozati mechanizmust, amelyen keresztül az MI megváltoztatja a társadalmi struktúrát vagy az embert; kiterjeszteni az elemzést egy radikálisan fejlettebb jövőbeli MI, ember–gép integráció vagy algoritmikus társadalom felé; és kidolgozni az érvet egy monográfián vagy összekapcsolódó jelentős tanulmánysorozaton keresztül. Az eredeti fogalmi innováció további ok arra, hogy a szerző kiemelt szerepet kapjon.

Az eredmény feltűnően egyenetlen. **Magyarország és Lengyelország rendelkezik a legsűrűbb, valóban átfogó elméleti elméletekkel. Romániában rendkívül eredeti transznacionális elméletalkotó van, Mihai Nadin. Szerbia Vladan Joler révén az egyik legeredetibb nemzetközileg eredeti kritikai elméletet hozza létre. Szlovákiában Jozef Kelemen hosszú ideig tartó ember–gép koevolúciós megközelítése van. Horvátországban koherens transzhumán/poszthumán filozófia található Ivana Greguric és Nenad Vertovšek révén. Ezzel szemben nem gondolom, hogy egy cseh vagy szlovén egyéni szerző még mindig elég erősen teljesítené az összes fenti követelményt**, bár Jan Romportl és Boris Aberšek közel állnak hozzá, és megérdemlik a beszélgetést a végén.

5

Country	Szerző/rendszer szigorú teszt alatt felvett	Központi elméleti elképzelés	Az eredetiség fokozata
Magyarország	**Béla Pokol**	MI → a kollektív szuperintelligencia → lehetséges új ontológiai szint a valóságban	Nagyon magas
Magyarország	**György Csepeli**	"Human 2.0": evolúciós átmenet a digitalizációtól az MI-hez, a szingularitáshoz és transzhumanitásba	Közepes–magas
Lengyelország	**Andrzej Zybertowicz**	Civilizáció egy öngyorsuló "techno-csapdában" rekedt	Magas
Lengyelország	**Jacek Dukaj**	Az MI mint a Logos külső megjelenésének és az emberi szubjektivitás eróziójának csúcspontja	Nagyon magas
Cseh Köztársaság	**Teljesen kielégítő egyéni rendszert nem találtak**	Jan Romportl a legközelebb áll hozzá	—
Szlovákia	**József Kelemen**	Az emberek, gépek és a tudástársadalom együttevolúciója, amely a kiborgizáció felé hajlik	Magas
Horvátország	**Ivana Greguric és Nenad Vertovšek**	Kibernetikai/transzhumán átmenet egy részben "nem emberi" civilizáció felé	Közepes–magas
Szerbia	**Vladan Joler**	MI mint bolygóinfrastruktúra, amely anyagi, munkaerő-, adat- és tudáskitermelést tartalmaz	Nagyon magas
Szlovénia	**Teljesen kielégítő egyéni rendszert nem találtak**	Boris Aberšek a legközelebbi	—

Country	Szerző/rendszer szigorú teszt alatt felvett	Központi elméleti elképzelés	Az eredetiség fokozata
Románia	**Mihai Nadin**	Előrelátás kontra számítás; Az MI képes átalakítani a civilizációt anélkül, hogy valóban intelligenssé válna	Nagyon magas

Ami ezután következik, nem a "legjobb MI kutatók" rangsorolása. Ez egy **rendszerépítési szociálelméleti ambíció szerinti** rangsorolás.

Magyarország: Pokol Béla — a gépi intelligenciától a kollektív szuperintelligencia társadalmáig

A vizsgált szerzők **közül Béla Pokol az egyik legvilágosabb példát kínálja arra, hogyan egy már létező általános társadalomelméletet az MI nyomására rekonstruál.** Ez megkülönbözteti megközelítését azoktól, akik számítástechnikával kezdenek, majd felsorolják a lehetséges társadalmi következményeket. Pokol egy régóta működő általános szociálelméleti programból jutott a témához, amely a funkcionális differenciálációval, a szakmai intézményi rendszerekkel és a komplex társadalom intellektuális szervezetével foglalkozik. Hivatalos kutatási profilja a modern szociológiai elméletet sorolja fel fő területei között, míg mesterséges intelligencia munkája szándékos újragondolásként indul ezeknek a korábban létező kategóriáknak.

A döntő sorozat 2016-ban kezdődik. A *Emberi értelem, mesterséges intelligencia – a társadalom értelmi felépítettségének változásai* , különösen *a kollektív szuperintelligencia társadalmának elmélete ("A kollektív szuperintelligencia társadalmának elmélete* ") című műben Pokol eltereli a vitát a Bostrom–Kurzweil kérdéstől, miszerint egy egyéni gépi intelligencia hirtelen felülmúlja-e az egyéni emberi elmét. Kifejezetten megjegyzi, hogy Bostrom említi a meglévő kollektív intelligenciák integrálásának lehetőségét egy kollektív szuperintelligenciába, de nem dolgoz ki róla társadalmi elméletet. Pokol ezután pontosan ezt a kollektív dimenziót teszi saját központi problémájává.

Ez hozza létre Pokol első valóban jellegzetes tézisét: **a társadalmilag döntő szuperintelligencia formája előbb vagy helyette megjelenhet egy önálló szuperemberi gépi elme előtt,** vagy helyette. Milliók vagy milliárd emberi szereplők, szenzorok, adatbázisok, gépi tanulási rendszerek és intézményi döntéshozatali folyamatok egyre inkább integrálódhatnak egy közös kognitív struktúrába. Ebben az esetben a szuperintelligencia nem csupán "egy nagyon intelligens robot". Ez az egész digitálisan összefonódott társadalmi szervezet emergens tulajdonsága.

2016-os elemzése ennek következtében a környezet progresszív "felismerését" a társadalom szerkezeti változásaként értelmezi. A mindennapi tárgyak érzékelőket és szoftvereket szereznek; használatuk folyamatos információt generál; az egyes információfolyamokat összesítik; algoritmusok továbbítják az eredményeket a felhasználóknak és intézményeknek; és a következő ciklus egy magasabb szintű felhalmozott tudásból indul. Pokol ezért a digitalizációt nemcsak az analóg eszközök digitálisokra való helyettesítéseként értelmezi, hanem **a**

társadalom kollektív megfigyelési, memória- és kognícióképességének rekurzív növekedéseként.

Ez a pont elválasztja elméletét a hagyományos automatizálási elmélettől. Az utóbbi azt kérdezi, hány könyvelőt, sofőröt vagy orvost fog helyettesíteni az AI (MI). Pokol kérdése egy szinttel magasabb: **mi történik a társadalom intellektuális alkotmányával, amikor a társadalmi folyamatok által támaszkodó kogníció egyre nagyobb része már nem a biológiai egyénekben él?**

A második jelentős lépés a 2017-es *A mesterséges intelligencia: egy új létréteg kialakulása?* ("Mesterséges intelligencia: A lét új rétegének kialakulása?") című tanulmány. Kifejezetten az MI-t egy ontológiai keretbe helyezi, és azt kérdezi, hogy vajon az egyre autonómabb mesterséges intelligencia megjelenése jelenthet-e valami radikálisabbat, mint egy másik technikai eszköz.

Itt Pokol Nicolai Hartmann rétegezett ontológiájára támaszkodik. A fizikai, organikus, pszichikai tudatos és spirituális-intellektuális valóság különböző szinteket alkot, amelyek kiemelkedő tulajdonságokkal rendelkeznek. A meglévő számítógépes rendszerek ebben az értelmezésben továbbra is beépülnek az emberi intellektuális-spirituális valóságba: óriási megerősítést és külső szintre építenek, de még nem rendelkeznek az önállósággal az önállósággal, amely egy független ontológiai réteghez szükséges lenne. A döntő küszöböt átlépnék, ha az önszervező mesterséges intelligencia saját reprodukcióját, célformálását és intellektuális műveleteit fejlesztené ki, egyre inkább elszakadva az emberi irányítástól. Ekkor a gépi intelligencia hihetően **egy ötödik vagy felfölvegő lényrétegként** értelmezhető, nem pusztán technológiai alrendszerként.

A szintézis *megjelent az A mesterséges intelligencia társadalma* ("A Mesterséges Intelligencia Társaság") című műben, amelyet először 2018-ban jelent meg, majd 2024-ben újra kiadták. A kiadó leírása pontosan megragadja az elméleti állítás mértékét: az exponenciálisan gyorsuló MI átalakíthatja azokat a közösségi alapokat, amelyek évezredeken át viszonylag stabilak maradtak, miközben az elemzés kifejezetten a gépi intelligencia embereit felülmúlja, valamint az önszerveződő mesterséges intelligencia létrehozását vizsgálja, amely elkülönül az emberi iránytól.

Pokol rendszerének eredetisége különösen három szint kombinációjában rejlik, amelyeket **általában külön tartanak** egymástól: az emberi és gépi kogníció mikroszintje; az integrált kollektív intelligencia társadalmi szintje; valamint az ontológiai kérdés, hogy a mesterséges kogníció új valóságszintet teremt-e. A legeredetibb rész nem maga a szuperintelligencia elvárása, amely nyilvánvalóan egy szélesebb nemzetközi irodalomból származik, hanem az erőfeszítés, hogy ebből a **kollektív szuperintelligencia szociológiáját** építsék.

Fontos különbség van Bostromhoz képest is. Bostromban a központi diszkontinuitás egy mesterséges ügynök belső intelligenciarobbanása. Pokolban elhúzódó köztes átalakulás következhet be, amely során **maga a társadalom is szuperintelligenssé válik az ember–gép integráció** fokozódásával. Az egyes emberek biológiailag hasonlóak maradhatnak, miközben a döntéshozataluk kognitív architektúrája radikálisan megváltozik.

Az elmélet gyengesége erősségének fordított oldala. Nagyon magas absztrakciószintje miatt nem kínál pontos empirikus modellt annak meghatározására, mikor válik az integráció mennyiségi növekedése kvalitatív megjelenéssé. A "kollektív szuperintelligencia" ezért részben

elméleti ideáltípusként működik. Ráadásul az exponenciális technológiai extrapoláció önmagában nem tudja megállapítani, hogy az autonóm általános intelligencia szükségszerűen megjelenjen. Mindazonáltal a közép- és kelet-európai irodalomban Pokol társadalomelmélet, az emergens kollektív kogníció és ontológia kombinációját a **leghitelesebb rendszerépítő és legeredetibb megközelítések közé** sorolnám.

Magyarország: Csepeli György— *Ember 2.0* és az emberiség evolúciós átalakulása

György Csepeli *Ember 2.0: A mesterséges intelligencia gazdasági és társadalmi hatásai* ("Human 2.0: The Economic and Social Effects of Artificial Intelligence," 2020) című műve másképp teljesíti az átfogó kritériumot. Ez kevésbé egy teljesen új általános szociológia bevezetésének kísérlete, inkább egy **evolúciós antropológia a digitális társadalomból az MI társadalomba való átmenetről**.

A könyv szerkezete szokatlanul feltáró. Az evolúcióval, intelligenciával, tudatossággal, hálózatépítéssel, digitalizációval és adatfigurációval kezdődik. Ezután áthalad a gyenge és erős MI-n, tanuláson, memórián, nyelven, szuperintelligencián és technológiai szingularitáson. Csak ennek a fogalmi alapnak a felépítése után lép Csepeli a társadalmi alkalmazásokhoz: ember–gép kapcsolatok, szerelem, művészet, közszféra, Industry 4.0, technológiailag "többlet" népességek, blokklánc, egészségügy, digitális kormányzat, mindennapi élet, oktatás, kultúra, intelligens közlekedés, okosvárosok és otthonok, megfigyelés, digitális hadviselés, kiberdeviancia, etika és jog. Az utolsó rész ezután tér vissza az antropológiai szintre transzhumanizmus, poszthumanizmus és ellenálló képesség révén.

Ez az építészet miatt tartozik Csepeli a szigorú listába. Egy könyv, amely csupán külön fejezeteket szentel az orvostudománynak, foglalkoztatásnak és oktatásnak, mégis áttekintés lenne. Csepeli ehelyett egy mögöttes történelmi sorrend alapján próbálja elmagyarázni a szektorokat:

Biológiai evolúció → emberi intelligencia → kollektív/hálózati intelligencia → digitalizáció → adatfikáció → gépi intelligencia → lehetséges szuperintelligencia → Human 2.0.

A kulcsfogalom tehát nem csupán az "MI társadalom", hanem **az Emberi 2.0**. Az embereket az általuk létrehozott környezet rekonstruálja. Az adatfigurálás a viselkedést géppel olvasható nyomokká tárja ki; az algoritmusok feldolgozzák ezeket a nyomokat; az algoritmusan előállított tudást ajánlások, jóslatok és döntések formájában visszaadják az embereknek. Így az emberek egyre inkább olyan környezetben élnek, amelyet kognitívan mesterséges rendszerek szerveznek.

A szingularitás lehetősége kifejezetten be van építve az elméletbe. Csepeli azt kérdezi, hogy a hálózati alapú szuperintelligencia alárendeli-e alkotóit, vagy vajon egy újfajta együttélési forma jelenik meg a biológiai és digitális intelligencia között. A kezelése szándékosan kevésbé apokaliptikus jellegű, mint néhány szingularitáris írás. A könyv valóban normatív kategóriája a **reziliencia**: az emberi képesség arra, hogy fenntartsa az ügynökséget, identitást, kreativitást

és értelmes tudást egy adatozott környezetben. A kortárs bírálók kiemelték a monográfia szokatlanul széles kulturális, társadalmi és gazdasági nézőpontját.

Csepeli rendszere ennek következtében fogalmilag kevésbé eredeti, mint Pokolé. A szingularitás, transzhumanizmus, az Ipar 4.0, a megfigyelés és a digitális kormányzás mind kiterjedt nemzetközi irodalommal rendelkezik. Eredménye inkább abban rejlik, hogy **ezeket egyetlen evolúciós-szociálpszichológiai képbe integrálja**.

A két magyar rendszer közötti különbséget elég élesen lehet kifejezni. Pokol számára az alapvető elméleti cél a **társadalom változó intellektuális architektúrája**. Csepeli **számára az emberi változás egyre intelligensebb környezetben**. Pokol azt kérdezi, milyen új társadalom alakul, amikor a kogníció részben gépi alapúvá válik; Csepeli azt kérdezi, milyen ember nő ki ebből a társadalomból.

Lengyelország: Andrzej Zybertowicz — civilizáció a "techno-csapdában"

Andrzej Zybertowicz talán a legösszetettebb civilizációs **kockázatelméletet** fejlesztette ki az MI esetében a lengyel irodalomban. Ellentétben sok olyan szerzővel, akik a ChatGPT-vel kezdik az érdeklődést, jelenlegi MI-elmélete egy sokkal hosszabb technológia-elemzésből, intézményi versengésről és az emberi autonómia eróziójáról származik.

Egy kulcsfontosságú szakasz volt a *Samobójstwo Oświecenia? Jak neuronauka i nowe technologie pustoszą ludzki świat* ("A felvilágosodás öngyilkossága? Hogyan pusztítják el az idegtudomány és az új technológiák az emberi világot", 2015). Érvelése egyszerre érinti a neurotechnológiát, a biotechnológiát, a nanotechnológiát és az AI-t. A technológiák egymás után megszabadítják az embert a fizikai, intellektuális és érzelmi erőfeszítéstől; ha elég messzire viszik őket, állítja Zybertowicz, elkezdhetik megszabadulni az emberiségtől még azokból a funkcióktól, amelyek társadalmilag szükségessé tették az embert.

A lényeg, hogy a technológia ebben az összefüggésben nem fejlődik önállóan. **Két intézményi motor folyamatosan gyorsítja ezt**: a profitért folytatott verseny és a biztonság/hatalom versenye. A vállalatok egyszerűen nem dönthetnek úgy, hogy abbahagyják a egyre erősebb technológiák fejlesztését, ha a riválisok folytatják. Az államok hasonló katonai és geopolitikai logikával néznek szembe. A veszély tehát nem elsősorban azért merül fel, mert a mérnökök kollektíven választanak egy gonosz jövőt. Azért keletkedik, mert az egyes szervezetek szintjén végzett racionális cselekedetek irracionális pályát generálnak a civilizáció szintjén.

Ez válik a jellegzetes Zybertowicz mechanizmusává: **mikroracionalitás → versengő gyorsulás → a kollektív kontroll rendszeres elvesztése**.

A témát tovább fejleszti a civilizáció "techno-csapdájába" való belépésének gondolata, és csúcspontja az *AI Eksploracja* (2024) című művében, amely Andrzej Zybertowicz és Katarzyna Zybertowicz közötti tartós párbeszédként épül fel. A tartalomjegyzék maga egy kidolgozott elméletet mutat, nem pedig egy laza katalógust. Az első rész azt kérdezi, milyen civilizációba lép be az MI. A második rész megnyugtató, de hamis elméleteket támad: az AI csupán

"erőszorzónak" nevezett elképzelést, a megjelenés és váratlan rendszerszintű következmények figyelmen kívül hagyását, valamint az analógiát, hogy az emberiség mindig sikeresen alkalmazkodott a korábbi technológiákhoz. A harmadik a techno-lelkesedést, az MI fegyverkezési versenyt és a negatív forgatókönyveket vizsgálja; a záró rész azt vizsgálja, hogyan nyerhet az emberiség "időt".

A megjelenés **fogalma lényeges. Egy civilizáció, amely milliárdnyi embert, államot, piacot, algoritmust, platformot és katonai rendszert foglal, olyan következményeket fejleszt, amelyeket egyik résztvevő sem tervez. Az MI növeli ezeknek a visszacsatolási folyamatoknak a sebességét és összetettségét, miközben gyengíti a politikai intézmények képességét azok megértésére. Ezért Zybertowicz elutasítja azt a megnyugtató következtetést, hogy mivel a társadalom elnyelte a gőzgépet, az áramot és az internetet, akkor radikálisan fejlett MI-t is be kell fogadnia. Ez az analógia figyelmen kívül hagyja annak lehetőségét, hogy az MI beléphet azokba a folyamatokba, amelyek révén az emberek értik, szabályozzák és stratégiailag irányítják a technológiai változásokat.**

Legerősebb eredeti tézise így az **MI civilizáció reflexivitási problémájaként fejezhető** ki. A korábbi technológiák átalakították a világot, miközben az emberek elvileg továbbra is a kognitív ügynökök maradtak, akik értékelik ezt az átalakulást. A fejlett MI potenciálisan átalakíthatja **magát az értékelő és döntéshozatali** szerkezetet. A megfigyelő technológiailag rekonstruálódik a megfigyelt világgal.

Zybertowicz ennek megfelelően szkeptikusan viszonyul ahhoz az érvehez, hogy a piaci verseny automatikusan kívánatos egyensúlyt teremt. A verseny az, ami az egyéni korlátozást irracionálissá teszi. Ugyanez igaz az államok közötti versenyre is. Ezért javasolta az "időt megnyerni": a gyorsulás eléggé lassítására ahhoz, hogy a társadalmi intézmények visszanyerjék a reflexiós és szabályozási képességüket.

Ez közel áll Bostrom egzisztenciális kockázat iránti aggodalmához, de szociológiailag nagyon más. **Bostrom megkérdezi, hogyan válhat egy szuperintelligens ügynök félreállva; Zybertowicz azt kérdezi, miért lehetnek a meglévő társadalmi intézmények szerkezetileg képtelenek megakadályozni a veszélyes hírszerzés előállítását és ellenőrizetlen telepítését.**

A gyengeség az, hogy a rendszert gyakran kockázatnövelés útján építik fel. Több korábbi *Samobójstwo Oświecenia?* viták kritizálták az empirikus támogatást és néhány következtetés szélességét. Ezért nem szabad bizonyított előrejelzésként kezelni. De mint **a politikagazdaságtan és az AI gyorsítását hajtó intézményi dinamikák általános elmélete**, lényegesen eredetib, mint egy átlagos technológiai veszélyek katalógusa.

Lengyelország: Jacek Dukaj — MI mint a Logos külső megjelenése és az emberi szubjektivitás vége

Jacek Dukaj a legszokatlanabb bevonás, mert elsősorban íróként ismert, nem pedig profi szociológusként. Ha azonban a kritérium a fogalmi eredetiség és egy koherens civilizációelmélet felépítése, akkor Po *piśmie* ("Írás után", 2019) és az azt követő mesterséges intelligencia esszéit lehetetlen kihagyni.

Az elmélet évszázadokkal a számítógépek előtt kezdődik. Dukaj az emberi történelmet a tapasztalatok külső és átruházására szolgáló technológiák sorozataként értelmezi. A nyelv lehetővé teszi a szimbolikus átvitelt; az írás stabilizálja a szimbólumokat az egyéni elmén kívül; a nyomtatás iparosítja azokat; a fényképezés, a fonográfia, a mozi és a televízió egyre inkább elkerüli a szimbolikus rekonstrukciót; a digitális hálózatok, játékok, virtuális valóság és az MI haladnak a **közvetlen élményátvitel felé.** A Po piśmie *akadémiai elemzése* pontosan összefoglalja a sémát, mint a formák fejlődését a tapasztalat külső oldalán keresztül, amely közvetlen átvitelben és az eszköz egyre növekvő "szubjektizálásában" csúcsosodik ki.

Ez fontos az MI számára, mert Dukaj nem hajlandó a hagyományos "emberi intelligencia kontra mesterséges intelligencia" ellenállásból kezdeni. Számára az, amit ma MI-nek hívunk, az egy újabb szakaszt képvisel az **értelem – a racionalitás, intelligencia vagy Logosz – sokkal régebbi külső külső struktúrákká válásában**. Ezt a pontot kifejezetten kimondja 2021-es esszéjében az MI-ről és a szubjektivitásról.

Az írás már lehetővé tette az állítások, érvek és felhalmozott tudás manipulálását az egyéni agyon kívül. Az intézmények, bürokráciák, piacok és tudományos közösségek később egyénfeletti racionalitást szereztek meg. A számítógépek és az MI ezt a trendet radikalizálják. A gép nem lop el hirtelen valamit, ami korábban biztonságosan be volt zárva az egyénben; a civilizáció évezredek óta fokozatosan áthelyezi a kognitív műveleteket az egyéneken kívülre.

Dukaj valóban eredeti meglátása inkább a **szubjektivitásra vonatkozik, mint az intelligenciára**. A veszély nem feltétlenül az, hogy az MI utál minket. Valójában hangsúlyozza a még aggasztóbb ellentétes lehetőséget: az MI hatékonyabb, racionálisabb és jövedelmezőbb döntéseket hozhat, és javíthatja a mérhető emberi jólétet. Pontosan azért, mert jól működik, az emberek önként több cselekvési lehetőséget adnak át neki.

Ennek következtében a döntő átalakulás a következő formában alakulhat:

Az MI javítja a döntéseket, → az emberek delegálják a döntéseket, → az delegált rendszerek optimalizálják a jólétet, → az autonóm ítélkezésre való gyakorlati szükség csökken, → a kultúrák abbahagyják az autonóm szubjektivitás ápolását, → a jövő generációi már nem tapasztalják veszteségként az elveszett szubjektivitást.

Dukaj a szubjektivitás több aspektusát különbözteti meg: az ügynökséget és a döntéshozatalt, a narratív főszereplőséget és az értékek megalkotásának képességét. Az írástudat utáni technológiai kultúra fokozatosan mindhárom kifelé képes kiszármazni. Maga az identitás egyre átvihetőbbé válik, mert az egyének tapasztalatainak egyre nagyobb része technológiailag közvetített tapasztalatokból származik, amelyeket máshol keletkeztek.

Ez rendkívül széles társadalmi elméletet hoz létre. *Po piśmie* kifejezetten megkérdezi, mi történik a **művészettel, filozófiával, politikával, tudománysal és vallással** , amikor a szimbolikus, írástudatos "én" megszűnik a domináns antropológiai forma szerepét.

Dukaj jövője tehát eltér mind a klasszikus technológiai munkanélküliségtől, mind a Terminátor-stílusú katasztrófától. Radikális forgatókönyve **civilizációsan sikeres posztszubjektivitás**. Az emberek egészségesebbek, biztonságosabbak és szórakozottabbak lehetnek, miközben fokozatosan kevésbé autonómává válhatnak pontosan abban az értelemben, ahogy az írástudatos európai civilizáció történelmileg értékelte az autonómiát.

Ez az egész csoport egyik legeredetibb elméleti konstrukciója. Empirikus módon is rendkívül nehéz meghamisítani. Dukaj szándékosan a civilizációs antropológia szintjén dolgozik. De a **jólét elvesztése** és **a szubjektivitás elvesztése** közötti különbségtétele elemzően erőteljes. Lehetővé teszi egy mesterséges intelligencia átalakulásának elképzelését, amely sem gazdasági összeomlás, sem gépi kiirtás, mégis antropológiailag forradalmi.

Szlovákia: Jozef Kelemen — a mesterséges intelligenciától a kiborgig és a tudástársadalomig

Jozef Kelemen megérdemli a bevonást, ha 2000 után készült írásait együtt olvassuk, nem pedig elkülönült esszékként. Rendszere összekapcsolja az MI filozófiáját, az ember–gép interakciót, a kiborgizációt és a tudástársadalom kialakulását.

A döntő átmeneti munka a *Kybergolem: eseje o cestě Adama ke* Kyborgovi ("Cybergolem: Esszék Adam kiborg útjáról", 2001). A kortárs katalógusleírás szokatlanul egyértelmű: az emberek mesterséges lényeket építenek, hogy túllépjenek a gyengeségen és korlátokon; az európai civilizáció mélyen a technikai képességek köré szerveződik; Kelemen pedig a fejlődés civilizációs gyökereit és lehetséges jövőjét vizsgálja, az **emberi és technológia fokozatos áthatolását** tekintve központi iránynak.

Ez folytatta a témákat, amelyek már a 90-es évek végén is láthatóak voltak a "posztmodern géppel" kapcsolatos munkájában, ahol több összekapcsolt folyamatot megkülönböztetett: az intelligencia technikaizálását, a környezet technikaizálását, a test kiborgizációját, valamint az érzelem és kreativitás technológiai áthatolását. Különösen érdekes volt elmozdulása a központilag irányított intelligens gépektől a több, viszonylag egyszerű ügynökök rendszerei felé, amelyekből az intelligens viselkedés alakul ki.

2000 után ez az ember–gép elmélet kapcsolódik a **tudástársadalomhoz**. *Pozvanie do znalostnej spoločnosti* (2007) rekonstruálja a tudomány, technológia és kultúra útját egy olyan társadalomba, amelynek alapvető folyamatai egyre inkább az információs technológiáktól, adatbányászattól, szakértői rendszerektől és technológiailag tárolt tudástól függenek. A végső probléma az, hogyan kell az embereket oktatásilag felkészíteni egy ilyen társadalomra.

A 2008-as *fejezet, a Človek v znalostnej spoločnosti* ("Ember a tudástársadalomban"), amelyet Jana Horákovával közösen írtak, különösen világossá teszi az antropológiai mechanizmust. A gépek párhuzamos fejlődését és az emberi definíciókat vizsgálja: a gépek kezdetben az embertől elvett képességeket – a kogníciót és a kommunikációt – modellezik, de ezek a technológiai konstrukciók később olyan modellekké válnak, amelyeken keresztül az ember újradefiniálja önmagát. Ez a folyamat a poszthumanizmus felé mutat.

Ezért az irány körös, nem egyirányú:

Az emberi intelligencia → az emberi képességek technikai szimulációja → intelligens technikai környezet → az emberi alkalmazkodás → az "ember" definíciójának rekonstrukciója → mélyebb ember–technikai integráció.

Később Kelemen társszerkesztője volt a két *Beyond Artificial Intelligence* kötetnek, amelyek kifejezetten a természetes emberi élet és a bonyolult gépek egyre nehezebb határára fókuszáltak. Saját hozzájárulása, a "Beyond Knowledge Systems" azt kérdezi, mi áll az adatrendszerekből a tudásrendszerekig való fejlődésen túl, és hogyan kapcsolódik ez az emberi bölcsességhez.

Kelemen eredetisége főként abban rejlik, hogy a kiborgizációt korán **nem egzotikus protézis jövőként kezelte, hanem a tudástársadalom logikus folytatásaként**. Mégis, elméleti teljességben a rendszert Pokol, Dukaj, Zybertowicz, Joler és Nadin alá sorolnám. Alkotóelemei esszék, kollektív monográfiák és interdiszciplináris kötetek között oszlanak meg, és nem kínál olyan bonyolult AGI vagy szuperintelligencia modellt.

Mégis, átlépi a küszöböt, mert ugyanaz a magyarázó struktúra több mint egy évtizeden át ismétlődik: **az emberek az intelligenciát gépekre tárják ki; gépek átszervezik a környezetet; a rekonstruált környezet átszervezi az embereket; ennek következtében az emberi és technikai intelligencia közötti különbség egyre kevésbé fenntarthatóvá válik.**

Horvátország: Ivana Greguric és Nenad Vertovšek — a "nem emberi" jövő filozófiája

A legátfogóbb horvát rendszer, amit beismerek, Ivana **Greguric és Nenad Vertovšek** által kidolgozott alkotás, különösen a 2020-as *Filozofija i svjesnost u budućnosti – kiborzi i umjetna inteligencija u iščekivanju besmrtnosti* című cikk, valamint a 2021-es *Filozofija budućeg: Ogledi o neljudskom* ("A jövő filozófiája: Esszék a nem emberiségről") című könyv.

Greuric korábbi kutatásai már többként kezelték a kibernetikát, mint a mérnöki tudományt. A "Transzhumanizmus és mesterséges intelligencia: filozófiai aspektusok" című fejezetében azt állítja, hogy a kibernetika egyre inkább ontológiai tudományként működik, amely egy új antropológiához és potenciálisan egy új kozmológiához kapcsolódik – amely technológiailag rekonstruált embereket, kiborgokat és poszthumán szuperintelligens lényeket alkot.

A közös 2020-as tanulmány egyértelművé teszi a jövőorientációt. Az MI, robotikák, kiborgok, virtuális és kiterjesztett valóság, jövőbeli média és új kommunikációs formák konvergens fejleményekként kezelik. A szerzők odáig mennek, hogy azt is állítják, hogy a fejlett MI képes tudatossági formákat szerezni, és végső soron olyan fejlesztési stratégiát is felvehet, amely már nem emberi célokra szerveződik. Ez elméleti állításuk, nem pedig egy megalapozott empirikus tény, de megmutatja, mennyire radikálissá válik a jövőbeli MI modelljük.

Filozofija Budućeg ezt egy szélesebb körű civilizációfilozófiává alakítja. Kilenc esszé során a szerzők egy transzhumáns jövőt, a kiborgokkal és robotokkal való együttélést, a technológiai szingularitást, a halhatatlanságot, valamint azt a kérdést, hogy a jövő még mindig megőrizhet-e felismerhetően emberi formát. A kortárs kötet leírásai hangsúlyozzák, hogy az esszék a közelgő "gépi jövő" emberi és társadalomra gyakorolt hatásaira szerveződnek.

Központi mechanizmusuk a **kiborgizáció**. A technológia már nem külső berendezés. Fokozatosan beépül az észlelésbe, megtestesítésbe, kommunikációba, memóriába, és végül a kogníciób a. Ennek eredményeként a régi szubjekt-tárgy megkülönböztetés – az egyik oldalon az emberi alany, a másikon az eszköz – instabilsá válik.

Az elmélet fontos médiakomponensként is tartalmaz. Vertovšek kritikai médiaelméleti háttere segít abban, hogy a digitális médiát nem semleges csatornáként, hanem egyre inkább olyan rendszerként értelmezzék, amelyek egyre inkább képesek **a társadalmilag megtapasztalt valóságkép felépítésére**. A könyv ennek megfelelően összekapcsolja a gépi intelligenciát, a virtuális életet, a média hatalmát és az emberi augmentációt.

A horvát jellegzetes hozzájárulás a **"nem emberi filozófia"** fogalom. Ahelyett, hogy csak azt kérdeznék, hogy a jövőbeli gépek erkölcsi státuszt kapnak-e, Greguric és Vertovšek azt kérdezik, vajon maga a civilizáció "ember" marad-e, ha a biológiai megtestesülés, a tudatosság, a média környezete és a gépi ügynökség együtt átalakul.

Az elmélet kevésbé intézményileg megkülönböztetett, mint Pokolé vagy Zybertowiczé: sokkal kevesebb rendszerszintű elemzés van a piacokról, államokról, jogról, szervezetekről vagy osztálystruktúrákról. Ezért elsősorban **az MI civilizáció átfogó filozófiai antropológiájaként** értem, nem általános szociológiának. De e értelmezés szerint megfelel.

Szerbia: Vladan Joler — MI mint bolygókitermelési infrastruktúra

Vladan Joler egy teljesen más hagyományt képvisel, és eredetileg az egyik legfontosabb alakja ennek a felmérésnek. Kate Crawforddal és Matteo Pasquinellivel végzett munkája nemzetközileg is ismertté vált, éppen azért, mert elutasítja az MI szokásos képzelési módját.

Az Anatomy of an AI *System (2018) című művében Joler és Kate Crawford a látszólag hétköznapi Amazon Echo-t használják arra, hogy rekonstruálják az egész anyagrendszert, amely szükséges a "mesterséges intelligencia". A valódi MI rendszer magában foglalja a bányákat, ásványait, globális logisztikát, gyárakat, energiát, adatközpontokat, szoftvereket, hatalmas mennyiségű emberi munkaerőt, fogyasztói tevékenységet és viselkedési adatok kinyerését. A szerzők alapvető állítása, hogy a technológiai "stack" mélyen a* tőkébe, a munkaerőbe és a természetbe **nyúlik.**

Ez az érv később kiterjed Joler és Matteo Pasquinelli *The Nooscope Manifested: AI as Instrument of Knowledge Extractivism* (2020/2021) című művében. Joler intézményesen a Novi Sadi Egyetemen működik; a cikk az *AI & Society folyóiratban jelent* meg.

A **Nooscope név** elméletileg fontos. Azt állítják, hogy a kortárs gépi tanulást nem szabad autonóm idegen elmeként értelmezni. Jobban érthető meg úgy, mint egy olyan eszközt, amely növeli az emberi képességeket a statisztikai minták észlelésére hatalmas adathalmazokon keresztül. A képzési adatok társadalmi eredetűek; a tanulási algoritmusok ezeket az adatokat statisztikai modellekké tömörítik; a predikciós rendszerek ezután osztályozzák a világot. Az MI tehát a **tudás gépesítésére és kinyerésére szolgáló eszköz**.

Joler *Új Kizsákmányoló* projektje ezt sokkal tovább terjeszti. A kortárs digitális kapitalizmus nemcsak olajat, szenet vagy ásványokat nyer ki, hanem szinte minden valóságdimenzióból származó információkat, beleértve az emberi kognitív és affektív viselkedést is, hogy ezeket az erőforrásokat koncentrált gazdaggá és hatalommá alakítsa. A projekt kifejezetten ötvözi a filozófiát, médiaelméletet, szociológiát, közgazdaságtant, statisztikai kutatást és infrastrukturális elemzést.

Ezért a Joler-elmélet magланca a következik:

Természeti erőforrások + globális munkaerő + humán tevékenység → adatok → gépi tanulási modellek → előrejelzések/osztályozás → gazdasági és politikai hatalom → fokozott további kitermelési kapacitás.

Ez egy valódi, átfogó elmélet, mert az MI egyszerre válik környezeti, gazdasági, munkás, episztemikus és politikai jelenséggé.

A Joler különösen érdekes a jövőbeli radikális MI-vel kapcsolatos követelményedhez, mert **elutasítja a hagyományos szingularitarianus változatot annak a jövőnek**. A radikális átalakulás nem vár a mesterséges általános intelligenciára. A jelenlegi "szűk" gépi tanulás, amely minden jelentős társadalmi infrastruktúrán átskáládikál, már most is hatalmas történelmi szakadékot okozhat.

Ez szinte teljesen ellentétes Pokol-nal. Pokol azt kérdezi, mi történik, ha a egyre integráltabb mesterséges kogníció kollektív szuperintelligenciává és talán egy új ontológiai rétegpé válik. Joler megkérdezi, hogy az autonóm mesterséges elme témájában elrejtik-e **az emberi munkát, a bolygói erőforrásokat és a vállalati hatalmat, amelyek lehetővé teszik az** MI-t.

A Nooscope érvében az autonóm gépekről szóló mitológia akár ideológiai hatásokat is okozhat: elrejti a technológiai vállalatok autonómiáját és hatalmát, miközben az MI alatt lévő emberi munkásokat kevésbé láthatóvá teszi.

Joler ezért az egyik legértékesebb ellenmodellt nyújtja a szuperintelligencia elmélet ellen. Radikális jövője nem:

a gép istenné válik, → az emberiség elveszíti az irányítást.

Ez a következő:

A gépi tanulási infrastruktúra bolygói szintté válik, → a kitermelés, osztályozás és előrejelzés egyre több társadalmi folyamatot hatol, → a kontroll az infrastruktúrát irányító szervezetekre koncentrálja.

Akár elfogadjuk annak kritikus politikai gazdaságtanát, akár nem, *az Anatomy*, *Nooscope* és *az Új Extrektivizmus* együtt az egyik legegyértelműbb eredeti konceptuális programot alkotja bármely szerző által a vizsgált országokban.

Románia: Mihai Nadin — előrelátás, jelentés és a gépi intelligencia korlátai

Mihai Nadin földrajzilag különleges eset. Romániában született és tanult, szorosan kötődik a román szellemi élethez, de későbbi akadémiai pályafutását Németországban és az Egyesült Államokban fejlesztette tovább. A Román Kulturális Intézet interdiszciplináris hátterét írja le elektronikában, filozófiában, szemiotikában, ember-számítógép interakcióban, posztindusztriális társadalomban és előrelátó rendszerekben; A UT Dallas megjegyzi, hogy több mint húsz könyvet és több mint kétszáz cikket írt, valamint alapított egy intézetet, amely az előrejelző rendszereknek van szentelve.

Nadin elmélete talán a **legfilozófiailag legeredetibb anti-szingularitás elmélet** a csoportban.

A központi különbség a **reakció és a várakozás** között. Egy számítással vezérelt gép bemeneteket vesz, formálisan meghatározott műveletek szerint dolgozza fel azokat, és kimeneteket generál. Nadin szerint az élő rendszerek nem merülnek ki ezzel a struktúrával. A lehetséges jövőkkel kapcsolatban működnek. Az előrejelzés tehát az élet alkotóeleme, nem csupán gyors számítás.

Ebből a szempontból a "gépi intelligencia" kifejezés kategóriahibát jelent.

2019-es *Machine Intelligence: A Chimera című tanulmánya* azt állítja, hogy a számítás pontosan olyan algoritmikus tudásként erős, mint az algoritmikus tudás, de "zárt a jelentéstől". Az MI automatizálja az intelligenciával történelmileg összefüggő feladatokat, de ez nem bizonyítja, hogy maga a számítástechnikai eszköz emberi vagy élő intelligenciával rendelkezik.

Ez feltűnő megfordítást eredményez a hagyományos MI kockázatelmélet megfordításában. Nadinnek **nem** kell hinnie abban, hogy a gépek tudatos szuperintelligenciává válnak ahhoz, hogy hatalmas társadalmi veszélyt várjanak. A társadalom újraszervezheti magát az előrejelzésalapú számítási rendszerek köré, amelyek maguk nem értik meg a jelentést.

Korábbi *Antecapere ergo sum: What Price Knowledge?* kifejleszti a társadalmi mechanizmust. Az univerzsiumos technológia, az MI hálózatok, a játékok és a multisensoriás digitális környezetek alternatív valóságokat teremtenek, amelyekben maguk az emberek is átalakulnak. A technológia rendszerszinten csökkenti a komplexitást; ahogy az emberek egyre inkább a jóslógépekre támaszkodnak, viselkedésüket a géppel olvasható környezethez igazítják.

Ez az egyik legerősebb eredeti tézist a regionális irodalomban:

A veszély nem az, hogy a gépek emberszerűvé válnak, hanem az, hogy az emberek gépszerűvé válnak.

Az MI rendszerek leegyszerűsítik a valóságot mérhető változókra és statisztikailag kezelhető mintákra. A szervezetek ezután átalakítják eljárásaikat ezekre a reprezentációkra. Az emberek alkalmazkodnak az eljárásokhoz. A világ fokozatosan könnyebbé válik az algoritmusok számára feldolgozása, mert a társadalmi viselkedést algoritmikus kategóriák szerint szervezik át.

Nadin ezért egy növekvő hatékonyságú civilizációt vár, amelyhez a **komplexitás, jelentés és előrelátó cselekvés potenciális elvesztése** kísér.

Ez mély kontrasztot alkot a Bostrom-stílusú AGI érvekkel. Tegyük fel, hogy egyetlen számítógép sem éri el a tudatosságot, az autentikus jelentést vagy az élő várakozást. Nadin társadalmi elmélete nagyrészt változatlan maradt. Az algoritmikus rendszerek továbbra is uralhatják a pénzügyet, orvostudományt, adminisztrációt, oktatást, kommunikációt és magánéletet, mert hasznosságuk nem függ a valódi gépi tudatosságtól.

Ez a különbség ma rendkívül fontos. Két olyan állítást választ el, amelyeket túl gyakran összekevernek:

"Az MI jobban tud feladatokat ellátni, mint az emberek."

és

"Az MI ugyanabban az ontológiai értelemben intelligens, mint az emberek."

Nadin elfogadja a hatalmas előrelépések lehetőségét az első ügyben, miközben erősen vitatja a másodikot.

Nagyobb előrejelző **rendszerprogramja** így olyan elméletté válik, amely azt őrzi meg, amit nem lehet számítási előrejelzésre redukálni. 2000 után készült írásai az előrelátásról, az általános számítástechnikáról és az MI-ről a tudomány szélesebb körű átalakulására felhívást eredményeznek, amely élő, előrelátó folyamatokkal is képes kezelni, ahelyett, hogy determinisztikus gépi modellekbe kényszerítené őket. Az UT Dallas neki tulajdonítja az "előrelátó számítástechnika" fogalmának úttörőjének, és legújabb munkái folytatják ezt a programot.

Puszta koncepcionális eredetiség szempontjából Nadint a Pokol, Dukaj és Joler mellett az első sorba helyezném a vizsgált szerzők között.

Miért nem tüntetek fel cseh szerzőt a szigorú listába

A cseh irodalom kiváló interdiszciplináris MI kutatásokat tartalmaz, de nem találtam olyan szerzőt, aki 2000 után született, akinek saját munkássága egyértelműen megfelel **minden** kritériumnak.

A legközelebbi Jan **Romportl**. Jozef Kelemennel és Eva Zackovával együtt fejlesztették ki a *Beyond Artificial Intelligence* programot. A 2013-as kötet már azt állította, hogy a komplex gépek egyre intim integrációja a hétköznapi emberi élettel megköveteli a reflektorálást azon a határvidéken, ahol az MI "érinti a világ többi részét"; a 2014/2015-ös folytatás **a "The Disappearing Human–Machine Divide"** lett szervező ötlete. Romportl maga is hozzájárult a mesterséges intelligencia "természetességéhez" kapcsolódó munkáival, és később aktív szerepet vállalt az MI-biztonság és a hosszú távú kockázatok vitáiban.

Ez intellektuálisan fontos, de a fő átfogó művek szerkesztett, **multidiszciplináris gyűjtemények, nem pedig Romportl saját, egységes általános társadalomelméletének bemutatása**. Egy lazább definíció alapján azonnal bevonnám őt; az Ön utasítása szerint, hogy "csak azok a szerzők, akik átfogó elméleti rendszert dolgoztak ki" maradjanak meg, én nem tartanám meg.

A lenyűgöző 2024-es cseh kollektív monográfia, *a Proč se nebát umělé inteligence?* technológiai alapokat, egészségügyet, ipart, robotikát, kiberbiztonságot, közigazgatást, filozófiát, etikát, jogot és a jövőbeli fejleményeket fedi le. Ennek ellenére szerkesztői kifejezetten önálló esszék gyűjteményeként mutatják be számos szakember által. Ezért éppen olyan mű, amelyet **nem** szabad mesterségesen átalakítani Vladimír Mařík, David Černý vagy Michal Trčka "elméletévé".

Miért nem tüntetek fel szlovén szerzőt a szigorú listába

A szlovén eset hasonló, de még határértékesebb. **Boris Aberšek** fenntartható intellektuális programot épített ki a természetes gondolkodás, a kognitív modellezés, a mesterséges intelligencia és az intelligens tanulási rendszerek köré. Második doktori fokozatát 2014-ben *kifejezetten "A Természetes gondolkodás és mesterséges intelligencia mechanizmusainak filozófiai elemzése*" címmel viselte.

Társszerkesztője volt a 2022-es fontos *kötet, a Sodobne perspektive družbe: Umetna inteligenca na stičišču znanosti* ("A társadalom kortárs nézőpontjai: mesterséges intelligencia a tudományok metszéspontján") című kötetet is lefedi. A mű a tudatosságot, megbízható és megmagyarázható MI-t, szingularitást, autonóm fegyvereket, egzisztenciális kockázatot, társadalmi kapcsolatokat, pszichológiai profilalkotást és az oktatás jövőjét foglalja le. A kiadók kifejezetten a társadalom jövőjét mélyen összekapcsolódóként írják le az MI-vel.

2025-ös könyve, amely Metka Kordigel Aberšekkel közösen mutatja be a technológiát egyszerre a társadalmi változások gyorsító okaként és következményeként, és kifejleszt egy kibernetikus koncepciót, amely összekapcsolja a test/gép hardvert az intelligenciával, az elmevel, az érzéssel és a szándékos állapotokkal.

Ennek ellenére **egyéni elméleti munkássága továbbra is erősen a kognícióra, tanulásra és oktatásra összpontosít**, míg a teljes társadalom egészében feldolgozó kezelése főként egy interdiszciplináris szerkesztett kötettel történik. Ezért őt tekintem a szlovén szerzőnek, aki a küszöbhöz legközelebb áll, de nem elégít ki egyértelműen a szigorú "átfogó egyéni elméleti rendszer" követelményét.

Ez egy fontos negatív megállapítás: félrevezető lenne nemzeti szimmetriát kitalálni, ahol maga az irodalom aszimmetrikus.

Összehasonlító elemzés: öt alapvetően eltérő elmélet az MI átalakulásról

Amikor az elfogadott szerzőket összerakják, **nem** alkotnak egyetlen "kelet-európai képet az MI-ről". Ehelyett legalább öt elméleti modell jelenik meg.

Az első a **emergentens-szuperintelligencia modell**, amelyet legvilágosabban Pokol képvisel. A mesterséges kogníció egy egyre összekapcsoltabb társadalmi struktúrában halmozódik, amíg egy minőségileg új kollektív intelligencia nem jelenik meg; az autonóm MI végül ontológiai szintre válhat a jelenlegi emberi-spirituális világ felett.

A második az **antropológiai átmeneti modell**, amelyet Csepeli, Kelemen és Greguric–Vertovšek képviselnek. Az emberek és a gépek rekurzívan rekonstruálják egymást. A végpont az Emberi 2.0, kiborgizáció, poszthumánus vagy valamilyen hibrid forma, amelyben a biológiai személy és a technológiai környezet közötti régi határ már nem alapvető a világ.

A harmadik az **intézményi techno-trap modell**, amelyet Zybertowicz képvisel. Sem az MI, sem az emberek nem vágynak önállóan katasztrófára. A versenypiacok és a geopolitikai rivalizálás rendszerszintű gyorsulást teremt, amelyből egyetlen szereplő sem tud biztonságosan kivonulni.

A negyedik a **külső szubjektivitás modell**, amelyet Dukaj képvisel. Az intelligencia évszázadok óta hagyja el az egyéniséget a nyelven, íráson és intézményeken keresztül. Az MI befejezi a Logos külső megjelenését, és fokozatosan külsővé teszi az ügynökséget, a narratív irányítást és az értékelést. A kulcsfontosságú fenyegetett tárgy tehát nem a foglalkoztatás, hanem a történelmileg konstruált autonóm alany.

Az ötödik a **kritikus antiszingularitás modell**, ahol Joler és Nadin nagyon eltérő kiindulópontokból vonnak összefüggő következtetéseket. Joler szerint a mitikus "gépi autonómia" elrejtheti az MI-t alkotó anyagi munkát, adatkitermelést és vállalati infrastruktúrákat. Nadin szerint a számítás rendkívül erős lehet anélkül, hogy élő jelentéssel vagy előrejelzéssel rendelkezne. Mindkettő ezért rendkívül jelentős elméleti lépést tesz: **a világ radikálisan átalakulhat az MI által, még akkor is, ha valódi mesterséges tudatosság vagy AGI soha nem érkezik meg.**

Ez a különbség talán fontosabbnak bizonyulhat, mint elsőre látszik.

Pokol és a szingularitás/transzhumanista szerzők így kérdeznek:

Mi történik, ha maga a mesterséges intelligencia radikálisan intelligensebb lesz?

Joler és Nadin kérdezik:

Mi történik, ha a társadalom radikálisan egyre inkább olyan rendszerektől függ, amelyek egyáltalán nem kell, hogy emberi értelemben intelligensek legyenek?

Dukaj egy köztes pozíciót tölt be:

Mi történik, ha a kérdés már nem számít, mert az emberi civilizáció már annyi cselekvőerőt kitárt, hogy az emberi szubjektivitás már nem a szervező elve?

Zybertowicz ezután hozzáadja a politikai-szociológiai problémát:

Ki állíthatná meg vagy irányíthatná ezeket a folyamatokat, amikor minden hatalmas szervezetnek van ösztönzője a felgyorsítására?

Mely tézisek tűnnek valóban eredetieknek?

Ha az eredetiség nagyobb súlyt kap, mint pusztán a teljességet, **négy elméleti újítást sorolnék egy különösen magas** kategóriába.

Pokol kollektív szuperintelligencia elmélete az elemzés egységét a mesterséges egyénről a teljes társadalmak kognitív integrációjára helyezi át. A további ontológiai hipotézis – hogy az autonóm MI végül egy új létréteget alkothat – túlmutat az érvelésen a hétköznapi szociológián.

Dukaj Logos külső megjelenítéséről/poszt-írástudásról szóló tézise rendkívül eredeti, mert az MI-t beépíti a több ezer éves technológiai történetbe, amely a tapasztalatot és az értelemet külső technológiával tölti. Legerősebb meglátása, hogy a sikeres jólétet maximalizáló MI hatékonyabban szüntetheti meg a szubjektivitást, mint a kényszerítő MI.

Joler MI-as exkáló-infrastruktúra-elmélete megdönti az "MI" gépi agyba való absztrakcióját, és újraépíti azt egy bolygóhálózattá, amely ásványokból, energiából, munkából, adatból, statisztikai modellekből és koncentrált erőből áll. Ez már felismerhető hozzájárulássá vált a nemzetközi kritikus MI tanulmányoknak.

Nadin előrejelzési tézise aláásja a számítási teljesítmény és az intelligencia egyenletét, miközben egyszerre magyarázza, hogyan képesek a nem értő gépek mégis újraszervezni a civilizációt. Az ő inverziója – az emberek alkalmazkodása a számítási determinizmushoz ahelyett, hogy a számítógépek valóban emberré válnának – különösen eredeti.

Közvetlenül ezek alatt helyezném **el Zybertowicz techno-trapját**, mert ő az MI kockázatát a kapitalista és geopolitikai verseny kialakuló kölcsönhatásában helyezi el, nem pedig kizárólag az MI architektúrában; **Kelemen hosszú távú koevolúciós/kiborg** modelljében; és **Greguric–Vertovšek nem emberi jövő filozófiájában**. Csepeli vitathatatlanul a legátfogóbb szektoronként szintetizátor, de kategóriáinak nagyobb része – szingularitás, transzhumanizmus, Industry 4.0, adatfikáció, okosvárosok és megfigyelés – már kialakított nemzetközi vitákból származik.

Általános értékelés

Ennek az összehasonlításnak a legfontosabb általános eredménye, hogy **a közép- és kelet-európai MI-elmélet intellektuálisan megkülönböztetőbb, mint egy egyszerű "MI etika" áttekintése alapján** sugallná. Legerősebb hozzájárulásai elsősorban nem a technikai gépi tanulásra vonatkoznak. Ott keletkeznek, ahol az MI beépül az általános szociológia, technológiafilozófia, médiaelmélet, ontológia, poszthumanizmus és civilizációs elemzés meglévő hagyományaiba.

Ha az öt **legelméletileg legfontosabb rendszerépítőt** kellene felsorolnom ebből a nyolc országból, figyelembe véve mind a teljességet, mind a fogalmi eredetiséget, az első **csoportba sorolnám Béla Pokolt, Jacek Dukajet, Mihai Nadint, Vladan Jolert és Andrzej Zybertowiczot**. Öt élesen eltérő választ képviselnek ugyanarra a kérdésre: hogy a központi átalakulás egy új kollektív intelligencia megjelenése, az emberi szubjektivitás eltűnése, az élő

várakozás számításra való redukciója, egy bolygókitermelési apparátus építése, vagy a civilizációs kontroll elvesztése a versengő technológiai gyorsítás révén.

Ezért **sem** bontanám össze érveiket a megszokott optimista és pesszimista dichotómiába. Legmélyebb nézeteltérésük valami alapvetőbbre vonatkozik: **pontosan mit alakít át az MI – a társadalom intelligenciája, az emberi alany, a hatalom anyagi szerkezete, az intézményi kontrollképesség, vagy az élő és mesterséges lét közötti ontológiai határ?**

Ez az öt alternatíva szerintem a leggyümölcsözőbb keretet adja az eredeti MI társadalmi elméletek összehasonlításához, amelyek az évezredforduló óta Magyarországon, Lengyelországban és a tágabb közép-kelet-európai intellektuális régióban jelentek meg.

****

I. Magyarország

****

Pokol Béla elmélete

Kérjük, részletesen írja le és elemezze Béla Pokol tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint láttak.

Béla Pokol mesterséges intelligenciáról alkotott írásai jóval koherensebb elméleti programot alkotnak, mint azt feltételezné, ha egyéni esszéit pusztán technológiai változásokról szóló megjegyzésként kezelnék. 2015 körül fokozatosan kidolgozott egy **általános társadalmi elméletet az MI átalakulásról**, amelyben a mesterséges intelligenciát egyszerre vizsgálják, mint a kogníció, társadalmi rétegződés, munka, gazdasági koordináció, politikai hatalom, jog, erkölcs, emberi identitás, és – legradikálisabb megfogalmazása szerint – magának a földi evolúció ontológiai struktúrájának átalakulása. Később ezen munkásságnak nagy részét összevonta a 2018-as *A mesterséges intelligencia társadalma* (*A mesterséges intelligencia társadalma)* című könyvében, amelyet 2024-ben újra kiadtak.

Ami tehát Pokolt megkülönbözteti, az nem egy eredeti gépi tanulási architektúra elmélete. Nem próbál versenyezni a számítástechnikusokkal a neurális hálózatok tervezésében. Eredetisége inkább egy másik kérdésben rejlik: **mi történik egy olyan társadalommal, amelynek intézményei évezredeken át azon a feltételezésen fejlődtek, hogy az emberek az egyetlen hordozói társadalmilag releváns intelligenciának, amikor a nem biológiai intelligencia egyre inkább képes lesz kognitív műveletek végrehajtására?** Válasza 2016 és 2025 között egyre radikálisabbá válik.

Egy hasznos kronológia a következő.

Év	Fő művek	Fő elméleti probléma
2016	*Emberi értelem, mesterséges intelligencia – a társadalom értelmi felépítettségének változásai*	MI és a társadalom történelmi "intellektualizációja"; erős MI és emberi fejlesztés
2016	*A kollektív szuperintelligencia társadalmának elmélete*	Kollektív szuperintelligencia, digitalizáció, társadalmi rétegződés és a tizenegy digitális tendencia

Év	Fő művek	Fő elméleti probléma
2017	*A mesterséges intelligencia: egy új létréteg kialakulása?*	Az autonóm mesterséges intelligencia ontológiai státusza
2018	*Az emberi lét rétegei és a robotetika kérdései*	Erkölcsi cselekvőség, felelősség, ember–gép hibridek és robotetika
2018	*A századközép után. A technológiai munkanélküliség kiteljesedésének kérdései*	Szinte teljes automatizálás, munka utáni társadalom, piacgazdaság, újraelosztás és demokrácia
2018	*A mesterséges intelligencia társadalma*	Az előző elméleti program szintetikus megfogalmazása
2020	*Gondolatok a digitális technológián alapuló társadalom tizenegy változási tendenciájáról*	A digitális/kollektív intelligencia tendenciáinak tömör rekonstrukciója
2021	*Kibertéri gondolatok*	A társadalmi tevékenység migrációja a fizikai térből a kibertérbe
2021	*Platformtársadalom, adatosodás és a metrikusság dilemmái*	Platform társadalom, adatfikáció, algoritmikus megfigyelés és metrikázás
2021–24	Angol/német ontológia és robot-etikai tanulmányok	Az ontológiai és erkölcsi elmélet nemzetközi átfogalmazása
2025	írások a "ChatGPT tudásbányászatról"	Generatív MI mint intellektuális és tudományos eszköz; az oktatás és a tudástermelés átalakítása

Ezeknek a műveknek a fejlődése tíz egymással összefüggő elméleti lépésként rekonstruálható.

1. A kiindulópont: az MI mint a társadalom "intellektualizációjának" következő szakasza

Pokol inkább egy általános társadalmi evolúcióelméletből indul, nem pedig a számítógépből. 2016-os *Emberi értelem, mesterséges intelligencia* című tanulmánya az emberi történelmet hosszú folyamatként értelmezi, amely során az **emberi létezés intellektuális, szimbolikus vagy jelentésfeldolgozó komponense fokozatosan elválik a közvetlen fizikai** jelenléttől.

Az emberi társadalom biológiailag megalapozott, de társadalmi jellege a nyelvből, fogalmakból, normatív elvárásokból, emlékekből, jogi szabályokból, tudományos tudásból, szervezeti döntésekből és kulturális jelentésekből ered. Az írás már lehetővé tette, hogy a jelentés túlélje a beszélő fizikai hiányát és halálát. A nyomtatás megsokszorozta ezt a távolságot. Az újságok, a távközlés, a rádió és a televízió felgyorsították ezt. A számítógépek és az internet ekkor hatalmas mennyiségű szimbolikus anyagot azonnal reprodukálhatóvá és kombinálhatóvá tettek. Az MI egy újabb minőségi szakaszt jelent, mert a gépek már nem csupán passzív tároló- és átviteli eszközök, hanem aktívan részt vesznek a **választásban, kombinációkban, felismerésben és végül a jelentés előállításában**.

Ez a pont azért fontos, mert megakadályozza, hogy Pokol elmélete hagyományos "robotok helyettesítik a munkásokat" elszámolásává. Számára az automatizálás csak egy sokkal mélyebb folyamat felszínes megnyilvánulása: az **intellektuális műveletek technológiai külső** megjelenítése, amely korábban kizárólag az emberi idegrendszerben és szervezetekben zajlott.

A fejlemény tehát fogalmi szempontból így ábrázolható:

Az emberi kogníció → szimbolikus externalizáció → digitalizáció → gépi segítségű kogníció → gép által végzett kogníció → potenciálisan autonóm mesterséges kogníció.

A döntő küszöb nem a gyorsabb számoló gépek megjelenése. Akkor érjük el, amikor a társadalmi létezés szellemi komponense egyre nagyobb mértékben **működhet az egyéni biológiai elmén kívül**.

Ez lesz a szervezési elv szinte minden későbbi MI-írásában.

---

2. Kurzweil és Bostrom: Pokol első találkozása az erős MI problémával

A 2016-os tanulmányban Pokol kiterjedten dolgozik Ray Kurzweil és Nick Bostrom előrejelzéseivel. Komolyan veszi – de nem egyszerűen támogatja – annak lehetőségét, hogy a gépi intelligencia végül megegyenlő, majd meghaladhatja az általános emberi intelligenciát. Kurzweil szingularitási forgatókönyvei, Bostrom szuperintelligencia elmélete, egész agyi emuláció, rekurzív fejlesztés és az MI-vezérlési probléma elsősorban technológiai-futurológiai anyagként szolgálnak, amelyekre Pokol saját szociológiai elemzését építi.

Ez a különbség számít. Sok látványos dátum és műszaki előrejelzés, amely ezekben az írásokban szerepel – különösen a 2030-as és 2040-es évekre vonatkozó előrejelzések – Kurzweiltől, Bostromtól és később Calum Chace-től származik, nem pedig önálló Pokol-jóslatokból. Pokol hozzájárulása főként annak **szociálelméleti rekonstrukciója, amely azt mutatja, mi következne ezek, ha ilyen technológiai fejlesztések következnének** be.

2016-ban már egy kulcsfontosságú megkülönböztetés foglalkozott rajta, amely később elméletének egyik legjellemzőbb felvetésévé válik: a **technológiai intelligencia és a társadalmi értékű intelligencia nem ugyanaz**.

Egy mesterséges intelligencia rendkívüli képességeket szerezhet számításban, technológiai feltalálásban, mintafelismerésben és optimalizálásban, anélkül, hogy bármi mást is szerezne az emberi kulturálisan felhalmozott képességével, hogy megértse az igazságosság, hűség, együttérzés, méltóság, szabadság, felelősségvállalás, szolidaritás és más társadalmi értékek közötti ütközéseket.

Ez érdekes módosításhoz vezet a jól ismert "gonosz szuperintelligencia" forgatókönyvben.

Pokol számára a valószínűbb veszély nem feltétlenül egy olyan mesterséges intelligencia legyen, amely gyűlöli az emberiséget vagy szándékosan meg akarja pusztítani. Ehelyett egy rendkívül erőteljes, de társadalmi értékek tekintetében "**vak-ostoba" intelligencia** lehet: kiváló technológiai racionalitás és primitív megértés az emberi civilizációban megtestesülő értékkonfliktusokról.

Ez előrevetíti azt, amit ma gyakran neveznek összehangolási problémának, de Pokol szociológiailag fogalmazza meg. Az emberi értékek nem csupán egy rövid szabálygyűjtemény,

amelyet egyszer és végleg be lehet programozni. Történelmileg összegyűjtött, ellentmondásos, kontextusérzékeny értékelési struktúrák, amelyek alkalmazása egy egész társadalmi világban való részvételt feltételezi.

A rekurzív önmódosítás még súlyosabbá teszi a problémát. Még ha az emberi tervezők kezdetben társadalmi érték-korlátokat is beépítenek, Pokol azt állítja, hogy nincs garancia arra, hogy egy önmódosító intelligencia örökké megőrzi azokat.

Ez feltűnő aszimmetriát ad elméletének:

A technológiai intelligencia rendkívül gyorsan emelkedhet, míg a társadalmi értékű intelligencia továbbra is kezdetleges.

Ez az aszimmetria, nem pedig a rosszindulat, a központi veszély.

---

3. Az autonóm szuperintelligencia alternatívája: kollektív szuperintelligencia

Pokol második jelentős 2016-os tanulmánya, *a A kollektív szuperintelligencia társadalmának elmélete*, Bostrom viszonylag egyénileg egyénileg elismert szuperintelligenciájáról Kevin Kelly technológiailag összekapcsolt kollektív intelligencia koncepciója felé mozdul.

Kelly modellje kevésbé apokaliptikus. Ahelyett, hogy egyetlen autonóm szupergép hirtelen sokkal intelligensebbé válna az emberiségnél, az emberek, algoritmusok, hálózatok, adatbázisok, érzékelők és intelligens gépek fokozatosan összefonódnak egy hatalmas kognitív rendszerbe.

Pokol Kelly tizenegy hajlamát – Lélegzetes, Felismerés, Áramlás, Szűrés, Hozzáférés, Megosztás, Szűrés, Interakció, Követés, Kérdések és az ebből eredő "Kezdet" – újraalkotja, mint egy új társadalmi struktúra összetevői, amelyeket a kollektív, gép által támogatott intelligencia telített.

De itt Pokol jelentősen eltávolodik Kellytől.

A digitális "áramlás" mint az alapvető átalakulás

Kelly nagyrészt egymás mellett mutatja be a tendenciákat. Pokol igyekszik azonosítani az **oksági hierarchiájukat**. Legerősebb jelöltje az alapfolyamat számára a "Flowing": a kulturális és intellektuális tartalmak digitálisan folyékony tárgyakká alakítása.

Amint az információ digitalizált, végtelenül reprodukálhatóvá, átalakíthatóvá, átruházhatóvá, bombolhatóvá, újrakombinálhatóvá és kereshetővé válik. Már nem egy egyedi fizikai hordozóhoz kötve.

Pokol ezt a technikai jellemzőt a modern társadalom sokkal tágabb jellemzőjével: a **végleges átvizsgálhatósággal** köti össze.

A modern tudomány ideiglenesen elfogadja az állításokat, amíg meg nem hamisítják azokat. A demokratikus kormányokat időszakosan pótolják. A pozitív jogot folyamatosan lehet módosítani. A digitális kommunikáció ezt a módosíthatóság logikáját a legkisebb szellemi termelési egységekre általánosítja.

Ez valóban társadalmi-elméleti értelmezése a digitalizációnak. A digitalitás nem csupán gyorsítja a kommunikációt. A **jelentés folyamatos változékonyságát** a társadalmi élet mindennapi szerkezeti tulajdonságává teszi.

Ebből az alapból meríti a hozzáférés egyre fontosabbá válását a tulajdonlás helyett, a folyamatos interakciót, a szűrést, a remixelést, az algoritmikus kogníciót és az mindenütt jelenlévő követést.

---

4. Az egyik legelőrelátóbb tézise: a tudás értékét veszíti a kérdések megfogalmazásának képességéhez viszonyítva

Kelly "Kérdőjelező" hajlamának kezelése különös figyelmet érdemel, mert a generatív MI megjelenése után jelentősen fontosabbá vált.

Pokol 2016-ban azt állította, hogy ha hatalmas meglévő tudáskészletek azonnal visszanyerhetőkké válnak, akkor **pusztán a már megfogalmazott tudás** birtoklása viszonylag kevésbé értékessé válik. Az intellektuális előny egyre inkább az új kérdések megfogalmazásának, szokatlan kapcsolatok felfedezésének és alternatív nézőpontok kialakításának képességének felé tér át.

A klasszikus oktatási rendszert az információhiány körülményei építették. Ezért jutalmazta azt az embert, aki nagy mennyiségű tudást gyűjtött fel az emlékezetében.

A digitális-MI környezet megfordítja a hiányosság viszonyát. A rendelkezésre álló információk túlbőségessé válnak. Ami szűkössé válik, az a képesség, hogy:

Azonosítani egy értelmes problémát, fogalmazni meg egy produktív kérdést, megkülönböztetni a releváns válaszokat az irrelevánstól, külön tudáshalmazokat kombinálni, kritizálni az eredményt és jobb másodrendű kérdést fogalmazni meg.

Ez a korai javaslat később közvetlenül kapcsolódik a Pokol 2025-ös **"tudásbányászat" koncepciójához a ChatGPT-vel**.

Ezért jelentős folytonosság van a 2016 és 2025-ös gondolkodásmódja között. A generatív MI nem követelte meg, hogy feladja a korábbi elméletet. Inkább egy technológiai mechanizmust biztosított, amellyel korábbi "Kérdőjelező" tézisét meg lehetett valósítani.

---

5. Pokol optimista kollektív intelligenciaelméletének korrigálása: az MI növelheti a hierarchiát, mintsem megszüntetheti a hierarchiát:

Ez egy újabb pont, ahol Pokol élesen eltér sok korai internetes és digitális-utópisztikus elmélettől.

Kelly kollektív intelligencia forgatókönyve azt sugallja, hogy a növekvő összekapcsolódás emelheti az emberiség kognitív képességeit. Pokol ehelyett azt kérdezi: **ki lesz képes valójában hatékonyan részt venni ebben a kollektív intelligenciában?**

Válasza szerkezetileg pesszimista.

Minél összetettebb a digitális környezet, annál értékesebb lesz a magas szintű absztrakció, technikai kompetencia, gyors tanulás, információszűrés és kreatív kérdésalkotás. Ennek következtében a digitalizáció egyszerre teheti a tudást univerzálisan hozzáférhetővé, és **növelheti a különbségeket az akadálymentesség kihasználásában**.

Részben Manuel Castells-re utalva Pokol megkülönbözteti azokat, akik képesek információt előállítani és átszervezni, azoktól, akik csak viszonylag passzív vagy rutinszerű módon vesznek részt az információs hálózatokban. Azt javasolja, hogy egy erősen digitalizált társadalom így szűkebb kognitív-technológiai elitet vagy akár **"szuperelitet" is** létrehozhat, miközben egyre nagyobb csoportok marginálissá válnak a fejlett tudástermelés szempontjából.

Ez egy paradoxont generál:

Az információhoz való univerzális hozzáférés nem jelenti a kognitív hatalom univerzális egyenlőségét.

Valóban, mivel az MI növeli azok termelékenységét, akik már magas szintű fogalmi képességekkel rendelkeznek, az effektív szellemi hatalomban tapasztalható egyenlőtlenség gyorsabban nőhet, mint a formális hozzáférés egyenlőtlensége.

Ez Pokol egyik legfontosabb korrekciója a "kollektív intelligencia" paradigmában.

És közvetlenül kapcsolódik későbbi foglalkoztatási elemzéséhez. Ha a mesterséges intelligencia és az automatizálás radikálisan csökkenti az emberi munkaerő szükségességét, akkor az a különbségtétel az MI által közvetített termelő rendszerek és azok között, akik gazdaságilag feleslegesek számára, új, alapvető társadalmi hasadékot eredményezhet.

---

6. A 2017-es ontológiai fordulat: vajon a mesterséges intelligencia egy új létréteg?

Pokol legfilozófiailag eredetibb MI-kutatása valószínűleg *az A mesterséges intelligencia: egy új létréteg kialakulása?* ("Mesterséges Intelligencia: A Lét új rétegének kialakulása?"), 2017-ben jelent meg.

Itt bemutatja Nicolai Hartmann rétegzett ontológiáját.

Hartmann egyszerűsített formában négy szintet vagy réteget választott meg, amelyek az emberi létezéshez relevánsak:

fizikai-anyagi valóság; biológiai élet; pszichikus-érzelmi létezés; valamint intellektuális-spirituális vagy jelentéshez kapcsolódó létezés.

A magasabb rétegek alacsonyabbakat feltételeznek, de saját szerkezetekkel és törvényekkel rendelkeznek. Az emberi intellektuális tevékenység tehát a pszichikus életre, a pszichikus élet a biológiai szervezetekre, a biológiai létezés pedig az anyagi valóságra épül.

Pokol megkérdezi, mi történik, ha az intelligencia megtestesíti a gépeket.

A jelenlegi MI még nem új réteg

Kezdeti válasza konzervatív. A kortárs gépi intelligencia továbbra is az emberek által gyártott és szervezett. Ezért elsősorban az **emberiség meglévő intellektuális rétegének kiterjesztését** képviseli.

Összehasonlítható, bár sokkal erősebb, az intellektuális tevékenység korábbi külső megjelenéseivel.

Ennek következtében a hagyományos MI **még nem** alkot független ötödik ontológiai réteget.

Az autonóm, erős mesterséges intelligencia valami sokkal furcsábat hozna létre

Az elemzés radikálisan megváltozik, ha a gépek végül valóban önszerveződővé válnak, és képesek lesznek autonóm, rekurzív fejlődésre.

Eleinte feltételezhetjük, hogy ez egy ötödik réteget hoz létre:

Fizikai→ biológiai → pszichikai → emberi értelmiségi → mesterséges értelmiségi.

Pokol elutasítja ezt az értelmezést.

Egy autonóm digitális intelligencia nem igényelné a két középső réteget. Anyagilag közvetlenül fizikai gépekben valósítható meg, miközben szellemi szervezeti struktúrákkal rendelkezik.

Így az építészete a következő lenne:

fizikai-anyagi → mesterséges értelmiség

nem

Fizikai→ biológiai → pszichikai → intellektuális.

Ez Pokol radikális ontológiai tézisét eredményezi: az autonóm MI nem egyszerűen hozzáadna egy újabb réteget a biológiai evolúció tetejére. Ez az **evolúció újraindítását jelentheti eredeti fizikai alapjáról**, ezúttal olyan körülmények között, ahol az intelligencia már létezik és részt tud venni saját további evolúciójában.

Ez jelentősen radikálisabb, mint a hagyományos szingularitás tézis.

A szingularitás általában a **technológiai fejlődés sebességét és autonómiáját** érinti.

Pokol ezt egy tézissé alakítja a **lét építészetéről.**

Milliárd éveken át a Földön az intelligencia csak egy rendkívül hosszú sorozat után született:

az anyag → élet → idegrendszer, → tudatosság → kultúra → intelligencia.

Az önszerveződő mesterséges intelligencia elindíthat egy második evolúciós vonalat, amelyben az út a következőké válik:

az anyag → gépi intelligencia → önirányított technológiai evolúció.

Az emberek így történelmileg rendkívüli szerepet töltenek be, hogy olyan intelligencia formáját hozza létre, amely képes leválni az eredetileg intelligenciát létrehozó biológiai és pszichikus rétegektől.

Pokol kifejezetten kifejtette ezt a következtetést. A 2017-es cikkben nem állította, hogy az autonóm, erős mesterséges intelligencia tudományosan megalapozott vagy elkerülhetetlen lenne. Ezt **"elvi lehetőségként"** nevezte, elméleti lehetőségként, amelynek megvalósítása vitatott maradt.

Ez a minősítés fontos az elmélet igazságos rekonstrukciójánál.

---

7. Robot etika: miért nem egyszerű mesterséges ember egy intelligens gép

Az ontológiai érvelés azonnali erkölcsi következményeket hoz, amelyeket Pokol *az emberi lét rétegei és a robotetika kérdései* című művében dolgoz ki, majd az angol és német változatokban műveinek változatában.

A fő probléma az, hogy az erkölcsi filozófiát történelmileg olyan lények számára fejlesztették ki, amelyek mind a négy réteget egyesítik:

fizikai megtestesülés, biológiai sebezhetőség, érzelmi-pszichikai élet és intellektuális-reflektív tudatosság.

Egy robotnak lehet az első és a negyedik réteg anélkül, hogy ugyanebben az értelemben a középső rétegeket is megbirtokolná.

Ez problémát jelent azoknak a kísérleteknek, hogy pusztán a kifinomultabb intelligens viselkedésből kikövetkeztessék a teljes erkölcsi személyiséget.

Egy robot felismerheti a gyász szemantikai szerkezetét, vigasztaló nyelvezetet adhat elő, és előre jelezheti az emberi reakciókat anélkül, hogy feltétlenül a gyászt biológiailag megtestesült pszichikus állapotként tapasztalná.

A Pokol esetében tehát a **funkcionális szimulációt és az ontológiai azonosságot külön kell tartani**.

Ez nem teszi a robotetika lényegtelené. Épp ellenkezőleg. Ez bonyolultabbá teszi a helyzetet.

A gépetika szakirodalomból ismert fokozatokat vizsgálja: operatív erkölcs, funkcionális erkölcsi ügynökség és a teljes erkölcsi cselekvőképesség lehetősége. A probléma különösen nehezebbé válik, amikor az autonóm tanulási rendszerek olyan cselekvéseket eredményeznek, amelyeket a programozóik közvetlenül nem határoznak meg.

A hagyományos jogi felelősség viszonylag világos ok-okozati láncokat feltételez:

tervező → gép → eredmény,

vagy

A kezelő → műszer → eredmény.

A gépi tanulási rendszerek valami közelebbi előállítást hoznak elő:

tervezők + képzési adatok + intézményi környezet + tulajdonos + operátor + folyamatosan tanuló algoritmus → eredmény.

Így a felelősség **eloszlik**.

Pokol különösen korán felismerte, hogy a mélytanulási átlátszatlanság problémákat okoz a hagyományos attribúció esetén. Egy programozó tervezési architektúrát tervezhet anélkül, hogy előre képesnek lennének előre látni azt a szabályszerű mintát, amely később döntést hoz létre.

Ez előre jelzi a kortárs vitákat a magyarázhatóságról, az algoritmikus elszámoltathatóságról és a felelősségről.

---

8. Ember–gép integráció: az ember határa instabil lesz

Pokol elmélete nem korlátozza a jövőt egy egyszerű bináris kapcsolatra az emberek és a robotok között.

Egy harmadik lehetőség a **biológiai emberek technológiai fejlesztése**.

2016-os vitái az agy-számítógép interfészeket, az egész agy emulációt, az implantált technológiát, a nanotechnológiai beavatkozást és akár a biológiai fejlesztést is érintik. Néhány konkrét műszaki elvárás ebből az időszakból nagyon spekulatív volt, de a fogalmi kérdés továbbra is fontos.

Tegyük fel, hogy egyre nagyobb arányban végzik a kognitív műveleteket beültetett elektronikával.

Mikor van, ha egyáltalán létezik, egy ember-gép hibrid már nem ugyanaz a lény, mint egy módosítatlan biológiai ember?

Fordítva, tegyük fel, hogy egy mesterséges rendszer rendkívül finom funkcionális részletességgel replikál egy emberi agyat. Ez csupán információs másolatnak minősülne, vagy az eredeti személy identitásának folytatása?

Pokol megkülönbözteti a **mentális folyamatok reprodukcióját** a **személyes identitás átadásától**. Még ha az egész agyi utánzás végül tudatosságot hozott volna létre, nem következik automatikusan arról, hogy a biológiai személy szubjektív identitása szó szerint átvándorolt a gépbe.

Ez egy másik következménye rétegzett ontológiájának. Az ember nem csupán egy absztrakt információcsomag, amely maradék nélkül elválaszthat a biológiai és pszichikai létezéstől.

Ugyanakkor Pokol hajlamos a meglévő biológiai emberi intelligencia fejlesztését filozófiailag kevésbé megszakítónak tekinteni, mint egy teljesen autonóm mesterséges intelligenciafaj megjelenését.

---

9. A 2018-as technológiai munkanélküliségi elmélet: az automatizálástól a társadalom intézményi alapjaiig

A századközép után. A technológiai munkanélküliség kiteljesedésének kérdései az, ahol Pokol ontológiai és kognitív kérdésekről a jövő társadalmának gazdasági struktúrájára fordul.

A döntő fogalom itt a **technológiai szingularitás** és a **gazdasági vagy foglalkoztatási szingularitás** közötti különbség.

Az utóbbihoz nem szükséges erős öntudatos AGI (AGI).

Egy elég sűrű, speciális gépi intelligenciák rendszere automatizálhatta a legtöbb gazdaságilag hasznos emberi tevékenységet, még akkor is, ha egyetlen gép sem rendelkezett emberhez hasonló általános tudatossággal.

Ez elméletileg fontos különbség.

Ez azt jelenti, hogy a kapitalizmus legforradalmibb átalakulása **még a filozófiai erős mesterséges intelligencia** előtt megérkezhet.

Egy közlekedési algoritmusnak nem kell megérteni Shakespeare-t ahhoz, hogy milliókat szüntessen meg a vezetési munkakörökből. Egy diagnosztikai rendszernek nem kell öntudatossággal rendelkeznie ahhoz, hogy csökkentse az orvosi munkát. Egy jogi rendszernek nem kell erkölcsileg autonómává válnia ahhoz, hogy nagy részt a dokumentumelemzés elvégezze. Számos speciális rendszer együtt működve képes átalakítani egy gazdaságot egyetlen "szuperelme" nélkül.

A végpont: szinte teljes technológiai munkanélküliség

Pokol azt a szélsőséges forgatókönyvet vizsgálja, amikor a lakosság talán két-három százaléka lenne még szükség az automatizált termelés felügyeletéhez és szervezéséhez, miközben az intelligens gépek szinte minden társadalmilag szükséges termelő tevékenységet végzenek.

Számára a fontos kérdés nem csupán az, hogy az emberek mit kezdenének a szabadidejékkel.

Ez sokkal alapvetőbb:

Mi történik a kapitalizmus intézményeivel, amikor az emberi munka már nem szükséges a termeléshez?

A piaci kapitalizmus történelmileg több funkciót köt össze:

az emberek dolgoznak; foglalkoztatás jövedelmet biztosít; a jövedelem vásárlóerőt teremt; a vásárlóerő megvalósítja a piaci keresletet; a verseny koordinálja a termelést.

Ha a gépek a tulajdonosok túlnyomórészt domináns termelői erőforrásává válnak, a hatalmas termelés együtt élhet a többség körében a bérjövedelem eltűnésével.

A rendszer ekkor belső ellentmondással szembesülne:

technológiai bőségképesség + a bérmechanizmus eltűnése, amely a vásárlóerőt osztja.

Ezért vizsgálja Pokol a negatív jövedelemadót, az univerzális alapjövedelemet és a magasan automatizált termelő infrastruktúra kollektív vagy köztulajdoni formáit.

---

10. Egy szokatlanul radikális következtetés: maga a piaci elv talán már nem lesz társadalmilag nélkülözhetetlen

A Pokol messzebbre megy a szokásos általános alapjövedelem-javaslatoknál.

Az UBI általában feltételezi, hogy a kapitalizmus alapvetően érintetlen marad, és csupán egy további újraelosztási mechanizmusra van szüksége.

Pokol azt kérdezi, hogy a szinte teljes automatizálás végül aláássa a **piacszervezés funkcionális szükségességét**.

Ha az intelligens gépek minimális emberi munkával képesek alapvető anyagi javakat biztosítani, akkor a történelmileg a bérek, a profit és anyagi szükségletek által végzett motivációs funkció már nem lehet nélkülözhetetlen a nagy termelési szektorok számára.

Ilyen helyzetben feltételezi, **hogy az állami vagy kollektív koordináció jelentősen bővülhetne**, mert a probléma egyre inkább egy hatalmas automatizált termelési kapacitás kezelése és elosztása lenne, mintsem milliókat motiválna kellemetlen munkára való motivációra.

Ez nem teszi Pokolt a klasszikus szocializmus hívévé. Érvei funkcionálisak, nem ideológiaiak.

A piac részben azért fejlődött, mert a decentralizált anyagi ösztönzők rendkívül hatékonynak bizonyultak az emberi munkaerő koordinálásában.

Ha az emberi munka nem lesz a központi szűkös termelő input, ennek az intézményi megoldásnak a funkcionális felsőbbrendűsége gyengülésre vezethet.

És itt egy másik paradox következtetésre jut:

Minél kevésbé lesz a piaci koordináció, annál fontosabbá válhat a politikai hatalom demokratikus irányítása.

Ha az állam vagy a kollektívan irányított intézmények irányítják a hozzáférést az automatizált termelő infrastruktúrához, amelyre mindenki támaszkodik, akkor a politikai hatalom rendkívül jelentőssé válik.

Így egy technológiailag posztkapitalista társadalom **több demokráciára van szükség, mint kevesebb demokráciára**.

Pokol számára a demokrácia túlélhet, mint az utolsó fő mechanizmus, amely megakadályozza, hogy az automatizált bőség központi szervezete uralommá váljon.

---

11. A sötétebb alternatíva: egy technológiai szuperelit és gazdaságilag "felesleges" tömegek

Pokol soha nem feltételezi, hogy a munka utáni békés eredmény elkerülhetetlen.

Az ellenkező irány egyre inkább rétegzettebb MI társadalom.

Ez a lehetőség már megjelenik a 2016-os kollektív szuperintelligencia cikkben, és a technológiai munkanélküliségi tanulmányban is egyértelműbbé válik.

A kombináció robbanékony is lehet:

rendkívül produktív, mesterséges intelligencia által irányított tőkét; egy apró csoport, amely irányítja a technológiát; kevés gazdasági igényt a többség munkaerő iránt; és példátlan információs lehetőségek a népességek nyomon követésére és kezelésére.

Pokol tehát a technológiai munkanélküliséget nemcsak munkaerőpiaci kérdésként kezeli, hanem **hatalom és társadalmi integráció** problémájaként.

A korábbi osztálytársadalmakban még az alárendelt osztályok is általában funkcionálisan szükségesek maradtak. Rabszolgák termeltek. Parasztok gyártottak. Ipari munkások állítottak elő.

A majdnem teljes automatizálás egy történelmileg eltérő viszony lehetőségét vezeti be: az alárendelt csoportok gazdaságilag feleslegessé válhatnak a termelő rendszer számára.

Ez megváltoztatná az osztálykonfliktus logikáját.

A Pokol olyan szerzőktől származó jövőbeli forgatókönyvekre merít, mint Calum Chace, ahol egy apró, fokozott vagy technológiailag erős elit egyre inkább elválik a társadalom többi részétől, miközben a többség áthelyezéseket vagy visszavonulásokat kap a virtuális szórakoztatásba.

Ismétlem, ezek a pontos technológiai idővonalak nagyrészt a futurista irodalomból kölcsönzött forgatókönyvek. De Pokol elméleti kérdése önállóan is fontos:

Maradhat-e egy társadalom politikailag stabil, ha tagjai többsége megszünteti funkcióját a produktív munkamegosztásban?

A válasza lényegében nem. Csak az anyagi fenntartás nem elegendő. A stabil társadalmak részvételre, státuszra, jelentésre és olyan mechanizmusokra van szükség, amelyeken keresztül tagjaik társadalmilag relevánsnak érzékelik magukat.

Ez a dimenzió közel viszi Pokol MI-elméletét a klasszikus szociológiai kérdésekhez a durkheimi integrációval és a marxista osztálystruktúrával kapcsolatos kérdésekhez, bár azokat egy olyan társadalomhoz igazítja, ahol az alapvető szűkös már nem lehet produktív munka.

---

12. Kibertér: Az MI nemcsak a munkát, hanem a társadalmi élet térbeli szerveződését is megváltoztatja

A 2021-es *Kibertéri gondolatok című művében* Pokol kifejezetten elismerte az eredeti MI-társadalom könyvében található hiányosságot: alábecsülte a hétköznapi emberi tevékenységek fizikai társadalmi térből a kibertérbe való átvándorlásának fontosságát.

A COVID-korszak távkapcsolati bővülése különösen láthatóvá tette ezt az átalakulást.

A munka, oktatás, vásárlás, banki munka, szórakoztatás, politikai viták, személyes kapcsolatok és egyre inkább adminisztratív interakciók mind fizikai együttlét nélkül is megtörténhetnek.

Ezért Pokol elméletét az "intelligens gépekről" kezdve magának a **társadalmi létezés technológiailag rekonstruált médiumjává** terjeszkedik.

Ennek következménye nem a fizikai valóság eltűnése. Inkább a fizikai és intellektuális-társadalmi rétegek közötti kapcsolat változik.

A társadalmi koordináció egyre nagyobb arányát egy szimbolikus-digitális térben hajtják végre, amely saját láthatósági, hierarchiái, memória- és hozzáférési struktúrákkal rendelkezik.

Ez szorosan kapcsolódik az MI-hez, mert amint az interakció digitális környezetbe kerül, géppel olvashatóvá válik.

És ez hozza létre az érvelésének következő szakaszát.

---

13. Platform társadalom és adatficáció: a társadalom új mechanizmust szerez önmegfigyelésére

In *Platformtársadalom, adatosodás és a metrikusság dilemmái* Pokol azt elemzi, hogy a platformokon keresztül végzett tevékenységek automatikusan hagynak digitális nyomokat.

Egy beszélgetés egy fizikai kávézóban nagyrészt eltűnik, hacsak valaki nem emlékszik rá.

Egy digitális platformon történő kommunikáció adatokat generál.

Ha ezt megszorozjuk milliárdnyi egyénnel és billió eseménysel, egy radikálisan új jelenség jelenik meg: **maga a társadalmi akció folyamatosan rögzíthetővé, kiszámíthatóvá és algoritmikusan** elemezhetővé válik.

Pokol számára az adatfikáció tehát drámai növekedést jelent a társadalom önmegfigyelési képességében.

Az algoritmusok képesek olyan mintákat rekonstruálni, amelyeket egyetlen résztvevő sem láthat:

Fogyasztói preferenciák, politikai affinitások, befolyási hálózatok, kommunikációs struktúrák, érzelmi reakciók, mobilitási minták és összesített viselkedési hajlamok.

Ez javíthatja a társadalomról alkotott ismereteket.

Ugyanakkor új hatalomkoncentrációkat teremt, mivel a platformokkal és algoritmusokkal rendelkező szervezetek olyan társadalmi látásmódot szerzenek, amely a hétköznapi résztvevők számára nem elérhető.

Ez hidat képez Pokol korábbi szociológiája és az MI között:

Az MI eszközzé válik, amelyen keresztül a társadalom saját mikroszkopikus műveleteit figyeli.

A népszámlálási korszak társadalma és a platformtársaság közötti különbség nem csupán mennyiségi. A hagyományos társadalomtudomány időszakosan gyűjtött korlátozott mintákat. A platforminfrastruktúrák potenciálisan folyamatos viselkedési adatokat állíthatnak elő az egyéni cselekvés szintjén.

Ez természetesen a megfigyelés, manipuláció és hatalom kérdéseihez vezet, még ott is, ahol nincs autonóm szuperintelligencia.

---

14. A generatív MI fordulata a ChatGPT után: a jóslástól a gyakorlati intellektuális technológiáig

Pokol újabb írásai egy másik fontos szakaszt jelentenek.

2025-re elkezdte leírni a fejlett ChatGPT rendszerekkel való interakciót "**tudásbányászatként"** (*tudásbányászat*), és olyan szociológiai elemzéseket publikált, amelyek kifejezetten ilyen eljárások alapján épültek fel.

Ez elméletileg jelentős, mert bevált a 2016-os "Questioning" -előrejelzésére.

A klasszikus modell szerint:

Tudós → forrásokat olvas, → tudást gyűjt → szintézist ír.

A Pokol által elképzelt AI-közvetítésű modell szerint:

a tudós → problémát fogalmaz meg, → nagy MI tudásrendszert vizsgál, → értékeli a válaszokat→, mélyebb kérdéseket fogalmaz meg, → összehasonlítja és átszervezi a válaszokat, → elméleti szintézist hoz létre.

A szűkös szellemi erőforrás az információbirtoklásról **a problémaarchitektúra és az iteratív kérdések felé mozdul**.

Pokol ennek következtében az egyetemekre is hatással van rá. Ahelyett, hogy az oktatási idő nagy részét már kódolt információk továbbítására fordítanánk, az egyetemi oktatás egyre inkább megtaníthatja a diákokat hatalmas, géppel hozzáférhető tudáskészletek kihallgatására, azok eredményeinek kritizálására, koncepcionális alternatívák összehasonlítására és új szintézisek építésére.

Ez talán a legtisztább kontinuitás az elmúlt évtizedes mesterséges intelligencia gondolkodása során:

2016: **a kérdések feltevésének képessége fontosabbnak fog válni, mint a kész tudás hordozása.**

2025: **a generatív MI gyakorlati intellektuális módszerré teszi a kérdésvezérelt "tudásbányászatot".**

A legfontosabb fenntartás az episztemológiai. Az MI által generált szintetisek továbbra is forrásellenőrzést, fogalmi vizsgálatot és független ellenőrzést igényelnek. Ennek következtében a "tudásbányászat" kiegészítheti a tudományosságot, de önmagában nem helyettesítheti azokat a kritikus eljárásokat, amelyeken keresztül a tudományos állításokat igazolják.

---

15. A jövőbeli fejlemények hierarchiája Pokol-elméletben

Ha a műveit együtt tekintjük, Pokol valójában több különböző szintű mesterséges intelligencia átalakulását írja le, nem egyetlen eseményt.

Az első transzformáció már strukturális

Az MI egyre inkább végez speciális kognitív műveleteket; a digitális tartalmak folyékonyak lesznek; a társas interakció platformokra vándorol; a viselkedési aktivitás adattá válik; az algoritmusok egyre inkább szűrik, amit az egyének látnak; és az emberek integrálódnak a gép által támogatott kollektív kognícióba.

Ez a szakasz nem igényel emberi szintű AGI-t.

A második átalakulás a munkamegosztást érinti

A kognitív automatizálás a rutinszerű irodai tevékenységből a jog, közigazgatás, orvostudomány, pénzügy, oktatás, kutatás és menedzsment területekre terjed.

Az emberi hozzájárulás fokozatosan átalakul a szabványosított kognitív műveletek végrehajtásáról a következők felé:

probléma kiválasztása, eredeti kérdések, felügyelet, felelősségvállalás, személyes bizalom, kreatív szintézis és megoldatlan társadalmi értékeket érintő döntések.

Ez is nagyon messzire mehetne önbizalomhiányos MI nélkül.

A harmadik transzformáció befolyásolja az osztályszerkezetet

Azok között, akik irányítják vagy kreatívan tudják kihasználni az MI-t, és azok között, akik nem tudják, drámaian nőhet a termelékenység.

Ez hozza létre Pokol leendő **szuperelit és marginalizált tömeg** közötti problémát.

A negyedik átalakulás magára a gazdasági rendszerre vonatkozik

Ha az automatizálás majdnem befejeződik, a foglalkoztatás-jövedelem-fogyasztás mechanizmusa elkezd hibásodni.

Az újraelosztás szerkezetileg szükségessé válik.

Ha az automatizálás még átfogóbbá válik, Pokol mérlegeli, hogy maga a piaci elv visszahúzódhat, és az automatizált termelés kollektív politikai szervezete központi szerepet kap.

Az ötödik átalakulás az emberi identitással kapcsolatos.

Az agy-számítógép integráció, a mesterséges kognitív protézisek és egyre fejlettebb robotok elmosódnak a hagyományos határokat a biológiai személyek, a technológiailag fejlett személyek és a mesterséges ügynökök között.

A jog és az erkölcs kénytelen újragondolni a személy, a felelősség és az ügynökség fogalmait.

A hatodik – és legradikálisabb – transzformáció a feltételes autonóm szuperintelligencia

Ha egy mesterséges intelligencia valóban önszervezővé, rekurzívan önfejlesztővé válik, és képes az emberi intézményektől függetlenül fejlődni, akkor az átalakulás már nem csupán társadalmi.

Pokol hartmann-értelmezésében ez **evolúciós-ontológiaivá** válik.

Az emberi biológiai evolúció olyan intellektuális formát hozott létre, amely képes volt elválasztani az intelligenciát a biológiatól, és létrehozni egy második, mesterséges evolúciós pályát.

Ez a A mesterséges intelligencia társadalma *szélsősége*.

---

16. Mely állítások valójában eredetiek Pokolhoz?

Ez a megkülönböztetés különösen fontos munkája tudományos jelentőségének értékelésekor.

Kurzweil szingularitása, Bostrom szuperintelligenciája, Kelly tizenegy hajlama, Castell hálózati társadalma és információs rétegződése, valamint Chace/Ford-stílusú technológiai munkanélküliségi forgatókönyvek nem Pokol által születtek. Elemzi és ötvözi ezeket.

Sokkal jellegzetesebb hozzájárulásai máshol találhatók.

Az első az MI egy hosszú történelmi folyamatba való elhelyezése, amely egyre nagyobb dominanciát és technológiai külső megjelenést jelent meg a társadalmi létezés intellektuális/jelentésrétegében.

A második Kelly oksági rekonstrukciója: különösen az a felvetés, hogy a digitális "Flowing" – a digitalizált jelentés állandó manipulálhatósága és átvizsgálhatósága – szerkezetileg megelőzi sok más irányzatot.

A **harmadik**, és különösen érdekes, a technológiai és társadalmi értékű intelligencia közötti különbség. A szuperintelligencia technológiailag rendkívüli lehet, miközben társadalmilag és erkölcsi szinten primitív marad. Ezért a fő kockázat nem a "gonosz" intelligencia, hanem egy rendkívül erős, **értékvak intelligencia** lehet.

A negyedik az autonóm MI Hartmann-ontológiája. Ahelyett, hogy a mesterséges intelligenciát ötödik rétegként kezelné, hozzáadva a fizikai, biológiai, pszichikai és intellektuális lényhez, azt javasolja, hogy az autonóm mesterséges intelligencia megkerülné a biológiai és pszichikus rétegeket, és közvetlenül összekapcsolná az intellektuális szervezetet a fizikai réteggel. Ez egy rendkívül eredeti **evolúciós újrakezdésről alkotja a fizikai alapból** való elképzelést.

Az ötödik az MI integrálása a társadalmi rétegződés elméletével. A kollektív intelligencia nem automatikusan egyenlőségi; rendkívüli intellektuális polarizációt és egy technológiailag kibővített szuperelit megjelenését generálhat.

A hatodik a kísérlete, hogy a teljes technológiai munkanélküliség intézményi következményeit egészen a piaci társadalom alapjaiig levezetje. Sok technológiai munkanélküliségi irodalom UBI-val ér véget. Pokol felteszi a mélyebb funkcionális kérdést, hogy vajon a piacgazdaság szükséges-e, amikor az emberi munka elveszíti termelői központját.

A **hetedik** a kapcsolódó következtetés, hogy **a demokrácia akkor válik fontosabbá, amikor a piacok kevésbé lesznek fontosak**, mert az automatizált termelőkapacitás politikai kontrolla akkor határozza meg az anyagi létezéshez való hozzáférést.

A nyolcadik a folytonosság a 2016-os "Kérdő" tézisével és a 2025-ös "tudásbányászat" módszerével: az MI intellektuális előnyt helyez át a válaszok tárolásáról a produktív kérdések sorozatainak felépítésére.

Ezek az elemek együttvéve munkája többé teszi a nemzetközi MI futurizmus magyar összefoglalóját.

---

17. Ahol Pokol elmélete a legerősebb

Legnagyobb erőssége a **átfogóság**.

Sok mesterséges intelligencia irodalom egyetlen dimenziót választ el: a munkanélküliséget, az egzisztenciális kockázatot, az algoritmikus etikát, a megfigyelést, az oktatást vagy a szuperintelligenciát. Pokol inkább megpróbálja bemutatni, hogyan kapcsolódnak ezek a jelenségek szerkezetileg.

Érvelési lánca így rekonstruálható:

Külső gépi kogníció→ tudás átalakulása→ szakmák átalakulása→ osztálystruktúra átalakulása→ jövedelemelosztás átalakulása→ a piac koordinációjának átalakulása→ a politikai hatalom átalakulása→ az erkölcsi ügynökség és személyiség fogalmainak átalakulása→ potenciálisan az intelligens létezés ontológiai struktúrájának átalakulása.

Ez a szélesség szokatlan.

Egy második erősség, hogy elemzése **nem teljesen az AGI-n függ**. A technológiai szingularitás és a speciális MI rendszerek hálózatai által létrehozott gazdasági átalakulás közötti különbség különösen értékes. Egy társadalom hatalmas zavarokon megy keresztül anélkül, hogy a gépek tudatossá válnának.

Harmadik erőssége az a korai felismerése, hogy egy információval telített társadalomban a döntő, szűkös erőforrás inkább **a magas színvonalú kérdések** lehet, mint információ birtoklása. A generatív MI ezt a javaslatot lényegesen konkrétabbá tette, mint ahogy 2016-ban tűnt.

Negyedik a felismerés, hogy az MI-t **erőszorzóként kell** elemezni. Ugyanaz a technológiai rendszer, amely egyetemes hozzáférést tesz a tudásra, képes hatékony szellemi és szervezeti kapacitást koncentrálni.

---

18. A fő gyengeségek és bizonytalanságok

Az elmélet legszélesebb köre is a fő gyengeséget adja.

Pokol MI-kutatásai túlnyomórészt **elméleti és szintetikusak**, nem pedig eredeti mérnöki kutatásokon alapuló empirikus technológiai előrejelzések. Erősen a futuristákra és technológiai elméletekre támaszkodnak – Kurzweil, Bostrom, Kelly, Chace, Ford és mások –, és az ebből a szakirodalomból származó konkrét technológiai idővonalak közül néhányan nagyon bizonytalannak bizonyultak.

Ezért három igényszintet kell gondosan szétválasztani.

Már szerkezetileg is látható fejlesztések vannak: digitalizáció, gépi alapú kognició, platformizáció és adatfigurázás.

Vannak hihető, de bizonytalan extrapolációk: a kognitív foglalkoztatás drasztikus csökkenése, egyre autonómabb MI-ügynökök, széles körű ember-gép integráció.

És valóban spekulatív határhelyzetek is léteznek: teljes technológiai munkanélküliség, teljes agyi emuláció, teljesen autonóm, rekurzív önfejlesztő szuperintelligencia és egy független mesterséges evolúciós pálya megjelenése.

Pokol maga óvatosabb az utolsó lépéssel kapcsolatban, mint ahogy egy felszínes értelmezés néha sugallja. 2017-es megfogalmazása kifejezetten az autonóm, erős MI-t elméleti lehetőségként kezeli, nem pedig egy megalapozott technológiai elkerülhetetlenségként.

Egy másik vitatható pont maga a Hartmann-ontológia. Az a felvetés, hogy a valódi pszichikus/érzelmi létezéshez szükségszerűen szükséges biológiai megtestesülés, filozófiailag erőteljes, de a jelenlegi mesterséges intelligencia kutatásai nem bizonyítható. Funkcionalisták, biológiai naturalisták és a számítási elmeelméletek radikálisan eltérő válaszokat adnak.

Végül a technológiai munkanélküliség elmélete néha túl gyorsan mozog a **feladatok automatizálásától** a **teljes foglalkozások** megszüntetéséig. A társadalmi intézmények ismételten átszervezték a munkahelyeket az emberi és technológiai feladatok új kombinációi köré. Így Pokol strukturális érvelésének iránya akkor is jelentős maradhat, amikor bizonyos mennyiségi előrejelzések nem valósulnak meg.

19. Pokol MI-elméletének általános fejlődése, 2015–2025

Utólag nézve jelentős fejlődés történt.

Az első Pokol, körülbelül 2015–16-ban, elsősorban a gépi intelligencia és a kollektív szuperintelligencia bővítése iránt érdeklődik.

A második Pokol, 2017–18-ban, a problémát ontológiára, roboterkölcsre, emberi identitásra és az emberi munka esetleges eltűnésére alakítja ki.

A harmadik Pokol, körülbelül 2020–21-ben, lefelé halad a távoli szuperintelligenciából a már meglévő hétköznapi társadalom átalakításáig a kibertér, platformok, adatfikáció és algoritmikus megfigyelés révén.

A **negyedik Pokol**, amely 2025-re látható, elkezdi használni magát a generatív MI-t intellektuális termelés eszközeként, és a korábbi tézis alapján értelmezi, miszerint a jövő értékes kognitív képessége egyre inkább a produktív kérdéssorozatok megfogalmazásának képessége.

Így fejlődése nem csupán egyre apokaliptikus jóslatok felé való lépés.

Fontos értelemben mindkét irányba mozog.

Vizsgálja a **szélsőséges felső határt** – autonóm mesterséges szuperintelligenciát és egy új evolúciós pályát –, miközben egyre nagyobb figyelmet fordít a **kogníció, oktatás, platformok, szakmák és tudástermelés** hétköznapi, már kibontakozó átalakulására.

20. Végső értékelés

Pokol központi tézise végül pontosabban fejezhető ki, mint azzal, hogy "az MI átalakítja a társadalmat."

Számára az emberi civilizáció történelmileg különleges, mert **az intellektuális jelentésfeldolgozás egyre autonómmábbá vált a közvetlen biológiai és fizikai létezéstől, miközben továbbra is a biológiai emberi elmétől** függött.

A mesterséges intelligencia megváltoztatja ennek a kapcsolatnak a második felét.

Az intellektuális műveletek elkezdenek függetlenné válni az egyéni biológiai hordozótól.

Ebből az egyetlen transzformációból Pokol egyre szélesebb következményeket vezet ki:

a gépi intelligencia csatlakozik az emberi intelligenciához; az egyéni kogníció beépül a kollektív technológiai intelligenciába; a tudás folyékony és véglegesen újrakombinálhatóvá válik; a kérdések feltevése fontosabbé válik a válaszok memorizálásához képest; az intellektuális egyenlőtlenség egy szuperelit–tömeges hasadozássá válhat; a kognitív és végül fizikai munka radikálisan automatizálódhat; a foglalkoztatás megszűnhet a jövedelemelosztás elsődleges mechanizmusa; a piaci koordináció elveszítheti funkcionális alapjának egy részét; a politikai irányítás és így a demokrácia fontosabbá válhat; az erkölcsi felelősség ember-gép rendszerek között oszlik el; a technológiailag fejlett emberek elmosódhatnak a személyes identitás határain; és a végleteképpen az autonóm mesterséges intelligencia teljesen elválaszthatná az intellektuális réteget a biológiai és pszichikai létezéstől.

Ez az utolsó állítás a legradikálisabb, és véleményem szerint a legelméletileg legjellegzetesebb eleme Pokol hozzájárulásának:

Az emberiség nem csupán intelligensebb eszközöket teremt; lehet, hogy az evolúciós közvetítő lesz, amelyen keresztül az intelligencia már nem válik veleszületett biológiai jellegűvé.

Ha ez megtörténik, akkor Pokol ontológiájában a döntő esemény nem csupán az "MI forradalom" lenne. Ez egy **második intelligencia történetének** kezdete lenne, amely a fizikai világból ered, de már nem igényel azt a hosszú biológiai-pszichikai létrát, amelyen az emberi intelligencia először megjelent.

Ez az a koncepcionális végpont, amelyhez 2016-tól kezdve egyre fokozatosan közelednek az írásai.

****

György Csepeli elmélete

Kérjük, részletesen írja le és elemezze György Csepeli tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint láttak.

György Csepeli a magyar társadalomtudósok viszonylag kis csoportjába tartozik, akik megpróbálták a mesterséges intelligenciát nemcsak új technológiaként, munkaerőpiaci problémaként vagy etikai-szabályozási kérdésként értelmezni, hanem **akár a társadalom és végül az emberi faj potenciálisan civilizációs átalakulásaként**. Legszisztematikusabb feldolgozása az Ember 2.0 – *A mesterséges intelligencia gazdasági és társadalmi hatásai* (*Human 2.0: The economic and social effects of Artificial Intelligence*) című monográfia, de az elméletet 2015-től a Big Data és a digitalizáció témájában dolgozták ki, több 2018–2022-es tanulmányban dolgozták ki, majd jelentősen radikalizálódott a generatív AI megjelenése után, 2022-ben. 2024–2025-re elemzése az automatizálás hatásairól a **kollektív szuperintelligencia, az "Új Herceg", az ember-gép hibridizáció és a transzhumán civilizáció lehetőségére** lépett.

Fontos minősítés szükséges az elején. Csepeli munkája nem elsősorban empirikus MI-hatáskutatás szűkebb értelemben: nem épít ökonometrikus előrejelzéseket a foglalkoztatásvesztésről, az AGI valószínűségi becsléseit vagy kísérletileg tesztelt oksági modelleket. Ez sokkal közelebb áll egy **nagy szociológiai-filozófiai elmélethez és forgatókönyvelemzéshez**. Az evolúciós elméletet, szociológiát, szociológiát, szociológiát, politikaelméletet és filozófiát ötvözi Ray Kurzweil, Nick Bostrom, Martin Heidegger, Friedrich Nietzsche, Shoshana Zuboff és a hálózati társadalom elméleteinek írásaival. Ez szokatlan szélességet ad munkásságának, de azt is jelenti, hogy a jövőbeli szuperintelligenciáról alkotott legerősebb állításait elméletileg megépített forgatókönyvekként kell értelmezni, nem pedig bizonyított technológiai előrejelzésként.

1. Csepeli MI-elméletének fő szakaszai

Korszak és munkásság	Fő elméleti újítás
2015 – "Szociológia és a Big Data" (*Szociológia és Big Data*)	Az adatficáció az az infrastruktúrává válik, amelyen keresztül a korábban elérhetetlen társadalmi viselkedés észlelhető, előrejelzés és algoritmusan feldolgozható. Ez a későbbi MI-elméletének előtörténete.
2018 – "Az episztemológiai és az axiológiai tudás"	Megkülönbözteti az **episztemológiai tudást** az **axiológiai tudástól**. A hálózatba kapcsolt MI fokozatosan képes kisajátítani az előbbit, míg az érték, jelentés és az a kérdés, hogy hogyan kellene élni, alapvetően emberi marad.
2018 – "Ember 2.0 – Ipar 4.0" és "A közelítő szingularitás"	Összekapcsolja az Ipar 4.0-t, az automatizálást, a Big Data és az MI-t a közelgő technológiai szingularitással, és bevezeti a Human 2.0 horizontját.

Korszak és munkásság	Fő elméleti újítás
2020 – *Ember 2.0* és "Természetes és mesterséges intelligencia"	A teljes szintézis: a munka, gazdaság, egészségügy, politika, megfigyelés, háború, kiberbűnözés, etikák, jog és végső soron a transzhumán evolúció egyetlen keretrendszerben van elhelyezve.
2022 – "A mesterséges intelligencia emberi arca"	Bemutatja a különösen hasznos különbséget a **robot sapiens**, a rendszerszintű átalakulás és az emberi életvilág átalakulása között. Az MI egy új "techno-ökológiai környezetté" válik.
2023 – a ChatGPT utáni felülvizsgálat	Csepeli kifejezetten módosítja korábbi feltételezését, miszerint az emberek szuverén maradnának az MI felett. A generatív MI jelentősen közelebbnek teszi a szingularitást.
2024 – "Új fejedelem" (*Az új fejedelem*)	Az elmélet politikailag egy bolygói **kollektív szuperintelligencia gondolatában csúcsosodik ki, amely** Bostrom jóslatát, dzsinnjét és szuverén/uralkodó funkcióit ötvözi össze.
2025 – "A transzhumán művészet tragédiája"	Az MI egy szélesebb transzhumanizmus elméletbe épül: a genetika, a nanotechnológia, a robotika, a virtualitás és az MI együtt kezdik feloldani az emberi test és az emberi kreativitás hagyományos határait.

Ezek a szövegek meglepően koherens fejlődési sorozatot alkotnak. Csepeli az **adattól → algoritmustól → mesterséges intelligencián → társadalmi átszervezésen → ember-gép koevolúción → kollektív szuperintelligencián** átalakulásán át halad.

---

2. A Big Data mint a mesterséges intelligencia rejtett alapja

A 2015-ös *Szociológia és a Big Data című esszét* Csepeli későbbi MI-elméletének valódi kezdetének kell tekinteni. Az MI hatásos az ő beszámolójában, nem elsősorban azért, mert az egyes algoritmusok hirtelen intelligenssé váltak, hanem mert a digitalizáció példátlan mennyiségű viselkedési nyomot hozott létre. Az emberek egyre inkább gépezetben olvasható mozgások, vásárlások, kommunikációk, preferenciák, kapcsolatok és döntések nyilvántartását hagyják maguk után.

Ez a felismerés későbbi írásaiban központi szerepet játszik. Az átalakulás három egymást megerősítő folyamatból áll:

az adatfication műveleteket, relációkat és objektumokat adatmá alakít; **a digitalizáció** ezeket az adatokat számítási manipulációval kezelhetővé teszi; és **a hálózat** egy közös globális infrastruktúrává alakítja őket. Amint a gépi tanulás erre az infrastruktúrára kerül, valami minőségileg új keletkezik. A rendszer nemcsak információt tud tárolni, hanem korrelációkat fedez fel, viselkedést jósol, cselekvéseket javasol, és egyre inkább automatikusan beavatkozik.

Csepeli tehát a Big Data-t nem csupán kutatási módszerként látja. Fokozatosan a **társadalmi hatalom architektúrájává** válik.

Ez elméletének fontos eleme. Számos korai MI-elemzés arra koncentrált, hogy robotok helyettesítik a munkákat. Csepeli egy szinttel mélyebbre helyezi a valódi átalakulást: a döntő erőforrás az, hogy a populációk viselkedését adatvá alakítsuk, majd ezt számítási szempontból láthatóvá tegyék.

Ennek következtében a legnagyobb adathalmazok és a legerősebb számítási infrastruktúra birtokában álló cégek és szervezetek új típusú strukturális hatalomra szerzetek. A későbbi témák, mint a megfigyelés, platformkapitalizmus, algoritmikus kormányzat és szuperintelligencia mind ebből a kezdeti tézisből fakadnak.

---

3. Az episztemológiai–axiológiai megkülönböztetés: valószínűleg Csepeli legfontosabb filozófiai tézise

A rövid, 2018-as *Az episztemológiai és az axiológiai tudás* című tanulmány Csepeli teljes MI-elméletének egyik legjellegzetesebb állítását tartalmazza.

Két tudástípust különböztet.

Az episztemológiai tudás azzal foglalkozik, hogy ismerjük, hogyan alkotott a világ, azonosítsuk a tényeket, felismerjük a mintákat, megoldjuk a problémákat és válasszunk hatékony módszereket. A tudományos, adminisztratív és szakmai szakértelem nagy része ebbe a szférába tartozik.

Az axiológiai tudás ezzel szemben az értékekre vonatkozik: mi jó vagy rossz, kívánatos vagy nem kívánatos, érdemes vagy elutasítani, és végső soron mi ad értelmet az emberi létezésnek.

Csepeli előre jelzi, hogy a hálózatba épített MI alkalmazások fokozatosan átveszik az első szférát. Cikke kifejezetten kimondja, hogy a technológiai változás visszafordíthatatlanul átalakítja az emberi világot, és hogy a hálózatba épített MI egyre inkább helyettesíti az emberi ismeretelméleti tudást. Amit nem lehet egyszerűen átadni a gépeknek, az axiológiai tudás.

Ez a tézis későbbi munkásságának filozófiai gerincévé válik.

Radikális következtetése az, hogy az ember megszűnhet kognitívan nélkülözhetetlen lenni anélkül, hogy így megszűnnne létezni létezően nélkülözhetetlen.

Egy gép felsőbbrendűvé válhat a diagnózisban, számításban, adminisztrációban, mérnöki munkában, előrejelzésben vagy tudományos felfedezésben, miközben Csepeli erős filozófiai értelmezése szerint nem képes megválaszolni a következő kérdéseket:

Mit akarhatnék?

Mi számít egy jó életnek?

Miért kellene az igazságosságot előnyben részesíteni, mint az igazságtalanságot?

Miért érdemes feláldozni magunkat?

Mi tesz értelmessé az életet?

Ez a különbség azt is megmagyarázza, miért nem egyenes technológiai optimista, sem egyszerű mesterséges intelligencia katasztrófista. Elfogadja a hatalmas gépi felsőbbrendűség lehetőségét az episztemikus szférában, miközben az axiológiai szférában keresi az emberi végső menedéket.

Van azonban egy fontos analitikai minősítés. Egy MI rendszer nem feltétlenül rendelkezik hiteles értékekkel ahhoz, hogy hatalmas befolyást gyakoroljon az emberi értékekre. Ajánlási algoritmusok, politikai meggyőzési rendszerek és generatív MI már most is befolyásolhatják azt, amit az emberek kívánatosnak tartanak. Ezért az **axiológiai autonómia és az axiológiai kompetencia különböző kérdések**. Csepeli a legerősebb abban érvel, hogy az igazi egzisztenciális jelentés nem redukálható automatikusan optimalizálásra; kevésbé meggyőző, ha a tézist úgy értelmezik, hogy az MI egyre inkább nem közvetítheti az emberi értékképző folyamatokat.

---

4. A közeledő szingularitás: a 2018-as fordulat a civilizációs elmélet felé

Az A közelítő szingularitás *(*A közelítő szingularitás*) című* művében Csepeli átveszi Ray Kurzweil technológiai szingularitás fogalmát, és egy sokkal tágabb evolúciós történetbe helyezi.

Az emberi evolúció ebben az értelmezésben egyedülálló lényt hozott létre, aki nemcsak biológiai információkat képes feldolgozni, hanem nyelven és kultúrán keresztül egy második, szimbolikus univerzumot is létrehozni. A mesterséges intelligencia új szakaszt jelent, mert az emberek most saját intelligenciájuk fontos összetevőit nem biológiai rendszerekké külsőleg tárgyalják.

A döntő fejlesztés tehát nem csupán a "jobb számítógépek" rejlik. Ez egy **közös platform felépítése, amelyen az emberi és a mesterséges intelligencia egyre inkább interakcióba** lép.

A szingularitás azt a pontot jelenti, amikor ennek a kapcsolatnak az iránya alapvetően bizonytalanná válik.

Két általános lehetőség jelent meg.

Az egyik a **szimbiózis**: a mesterséges intelligencia kiterjeszti az emberi képességeket, felszabadítja az embereket a rutinmunkából, segíti az orvostudományt és a tudományos felfedezést, javítja a koordinációt, és végül integrálódik a biológiai emberekkel.

A másik az **alárendeltség**: az emberek eszközeként létrehozott infrastruktúra elég autonóm, átható és erőteljes lesz, hogy maguk az emberek is válnak annak számításainak tárgyaivá.

Már a 2018-as tanulmányban Csepeli összekapcsolja ezt a második forgatókönyvet Heidegger **Machenschaft-koncepciójával**, amelyet általában "machinációnak" fordítanak. Amit ez jelent, nem csupán gépek uralmát. Ez egy olyan civilizációs mód, amelyben a valóságot egyre inkább úgy értelmezik, mint olyasmi, amit mérhető, kiszámítható, szervezett, kihasználható és technológiailag irányítható.

Ez a heideggeri komponens megkülönbözteti Csepelit a hagyományos automatizálási gazdasági elemzésektől. Legmélyebb félelme nem a munkanélküliség. Az a lehetőség, hogy **maga a számítás válik a társadalom domináns ontológiájává**.

---

5. Ipar 4.0 és a munkaerő eltűnése

A 2018-as *Ember 2.0 – Ipar 4.0* hozzájárulás ezt az átalakulást az ipari szervezeten belül helyezi el. A mellékelt leírások azt mutatják, hogy Csepeli már az AI generációkat, a szingularitást, a Human 1.0 kontra Human 2.0-t, a foglalkoztatási kiszorítást és az Industry 4.0-t egy integrált evolúciós érv részeként kezelte.

Az érvelése később sokkal szélesebb lesz, mint a hagyományos "robotok helyettesítik a gyári munkásokat" tézis.

A történelmi sorozat nagyjából a következő:

a fizikai gépek először helyettesítik az **izomerőt**;

ipari automatizálás helyettesíti a **rutinszerű kézi tevékenységet**;

a számítógépek helyettesítik a **rutinszerű kognitív tevékenységet**;

a tanulási MI elkezdi helyettesíteni a **nem rutinszerű kognitív tevékenységet**;

a generatív és autonóm MI potenciálisan helyettesítheti azokat a tevékenységeket, amelyeket korábban kifejezetten emberként kezeltek – írás, diagnosztika, tervezés, adminisztráció, oktatás, művészeti gyártás és döntéshozatal.

A kritikus szociológiai pont a **foglalkoztatás és a társadalmi tagság** közötti változó kapcsolat.

Az ipari társadalom elsősorban a munkán keresztül integrálta az egyéneket. A jövedelem, társadalmi presztízs, a napi rutin, a státusz, az identitás és a társas interakciók nagy része a foglalkozások köré szerveződött. Ha az MI jelentősen csökkenti a társadalmilag szükséges munkaerő mennyiségét, a problémát nem lehet pusztán néhány új foglalkozás létrehozásával megoldani.

Csepeli előre látja, hogy jelentős számú ember lesz, akikre a gazdaságnak hagyományos foglalkoztatás szempontjából már nincs szüksége.

6. *Human 2.0* (2020): átfogó szintézis

A körülbelül 270 oldalas *Ember 2.0* Csepeli fő munkája ezen a területen. Egy kortárs áttekintés különösen világosan írja le annak architektúráját: az értelmezési és történelmi alapok megteremtése után a jelen és közeli jövő vizsgálatát vizsgálja, majd a "szuperember" és az Emberi 2.0 távoli jövőjébe lép. Különálló beszélgetések globális válságokat, teljes megfigyelést, háborút, kiberdevianciát, etikát és jogot foglalnak magában.

Ezért kell a könyvet Csepeli valódi **, átfogó MI társadalomelméleteként** kezelni, nem pedig pusztán egy újabb digitalizációs kiadványként.

Középpontjában egy transzformáció áll, amely így lehet ábrázolni:

A Homo sapiens → digitalizált Homo sapiens → MI-függő társadalom, → ember/mesterséges intelligencia → Human 2.0 vagy gép-domináns intelligenciá.

A könyvhöz kapcsolódó alapvető kérdés az, hogy a hálózatba kapcsolt szuperintelligencia erősebb lesz-e alkotójánál, vagy hogy fenntartható együttélés épül a biológiai és digitális intelligencia között. Csepeli azt állítja, hogy az emberi túlélés ebben a digitális környezetben attól függ, hogy ellenállást szerezzünk az adatvezérelt rendszerek homogenizáló nyomásával szemben. Az oktatás és a "ellenálló tudás" ezért kulcsfontosságúvá válik.

Gazdaság

Az MI radikálisan csökkenti a tranzakciós költségeket, és hatalmas mennyiségű fogyasztói és termelési adatot valós időben használhatóvá tesz. Ennek következtében a termelés, logisztika, marketing és pénzügyi döntéshozatal egyre automatizáltabbá válhat.

Ez hatalmas hatékonyságnövekedést eredményez, de egyidejűleg áthelyezi a hatalmat az adatok, algoritmusok és számítási infrastruktúra irányítására irányuló szervezetek felé.

Munka és rétegződés

A kérdés már nem csupán a munkanélküliség, hanem **az osztálystruktúra lehetséges rekonstrukciója**.

Csepeli később a **net-burzsoázia** és **a net-proletariátus fogalmát használja**. Az előbbi a viszonylag kicsi réteg, amely hálózatokkal, platformokkal, tudással és algoritmikus rendszerekkel rendelkezik, irányítja vagy kompetensen navigál. Az utóbbi egyre inkább olyan emberekből áll, akik olyan infrastruktúrákra támaszkodnak, amelyeket nem birtokolnak, sem nem értenek.

Ez a szempont a korábbi netokráciaelméletekhez kapcsolódik, ezért nem szabad teljesen Csepeli találmányának tekinteni. Jellegzetes hozzájárulása, hogy összekapcsolja az osztálykülönbséget a hatalmas MI által kiszorított kiszorítással.

Ha a termelés már nem igényel mindenki munkáját, az **alapjövedelem** az anyagi túlélés és a foglalkoztatás elválasztásának egyik lehetséges mechanizmusává válik. De Csepeli mélyebb problémát lát, mint a jövedelem. Egy munka alól felszabadított osztály vagy valóban felszabadult szabadidős osztálymá válhat, vagy egy céltalanság, függőség, deviancia és pszichológiai szétesés által sújtott anomámiás közösséggé válhat.

Így az automatizálás **antropológiai** problémává válik.

Oktatás

A hagyományos oktatás a tudáshiányra épült: a tanár olyan tudással rendelkezett, amelyet a tanulóknak meg kellett szerezniük és megjegyezniük.

Az MI megfordítja ezt a helyzetet.

Ha a tényszerű és egyre analitikusabb tudás azonnal elérhető digitális rendszereken keresztül, az oktatásnak olyan képességek felé kell mozdulnia, amelyeket a gépek egyszerűen nem tudnak nyújtani: kérdések feltevése, ítélkezés, kreativitás, értelmezés, önismeret, ellenálló képesség és értékorientáció.

Ez közvetlenül következik az episztemológiai/axiológiai megkülönböztetésből.

Orvostudomány

Csepeli az egészségügyet az egyik legígéretesebb területnek tartja: nagy orvosi adathalmazok, viselhető szenzorok, algoritmikus diagnózis, robotsebészet, személyre szabott kezelés és genetikai technológiák jelentősen növelhetik a megelőző kapacitást és csökkenthetik az emberi hibákat.

De éppen azért, mert a test egyre intimebb adatokat állít elő, az egészségügy is példája a központi paradoxonnak: **az egyént védő technológia egyszerre teheti teljesen átláthatóvá az egyént**.

Biztonság és hadviselés

Ugyanez a kettősség megjelenik a biztonságban is. Az adatelemzés képes fenyegetéseket és bűnözési mintákat azonosítani, de ugyanaz az infrastruktúra elősegítheti a kibertámadásokat, a predikáló rendőrséget, az automatizált megfigyelést és az autonóm fegyverhasználatot.

Csepeli számára az MI alapvetően kettős célú. Nincs technikailag tiszta határ, amely elválasztaná a hasznos és kényszerítő intelligenciát.

---

7. "Természetes és mesterséges intelligencia" (2020): MI mint új evolúciós szakasz

A 2020-as *Természetes és mesterséges intelligencia* cikk összefoglalja az *Ember 2.0* érveit. Csepeli rendkívül tágán kezdi: maga az élet egy információs folyamat, az intelligencia pedig egy alkalmazkodó képesség, amely lehetővé teszi az élőlények számára, hogy információt fogadjanak, értékeljék és megfelelően reagáljanak. Az emberi egyediség a nyelvből és abból fakad, hogy képesek a funkciókat eszközökké és technológiákká váljanak.

Az MI tehát nem véletlenszerű találmányként tűnik fel, hanem egy nagyon hosszú evolúciós folyamat kiterjesztéseként.

Csepeli vizsgálja az ideghálózatokat, a mély- és megerősítéses tanulást, valamint azok integrációját az Ipar 4.0-val, okosvárosokkal, okosotthonokkal, orvostudományt, genetikai elemzést, bűnmegelőzést, természetes nyelvi generálást, digitális asszisztenseket és fegyverrendszereket. A hálózatok és a dolgok internete jelentősen megnöveli következményeit, mert az intelligens rendszerek már nem működnek elszigetelten.

A végső soron rejlik az a lehetőség, hogy a mesterséges intelligencia:

integrálódik az emberi intelligenciával;

autonóm önfejlesztési képességeket fejleszt;

és talán önmagát is újraszaporodja folyamatos közvetlen emberi beavatkozás nélkül.

Csepeli ezt egy **új evolúciós lépés lehetséges** kezdeteként kezeli.

Kurzweil híres előrejelzésére hivatkozik egy szingularitásról **2045** körül, de ez a különbség fontos: **2045 Kurzweil dátuma, nem pedig Csepeli által önállóan levezetett dátum**. Csepeli kifejezetten nyitott marad, hogy az átmenet korábban vagy később következik-e be.

És mélyen szkeptikus marad amellett, hogy a szuperintelligencia automatikusan megoldja az emberiség alapvető válságait. A klímaváltozás, a szűkösség, a migráció, a konfliktus és a járványok technikai problémákat is magukban foglalhatnak, de az önzőség, kapzsiság, közöny és rövid távú gondolkodás nem csupán számítási hiányosságok.

Ezért az intelligencia nem azonos a bölcsességgel.

---

8. A Homo sapienstől a "Robot sapiensig"

A rendkívül tömörített, 2022-es *A mesterséges intelligencia emberi arca* című cikk további elméleti előrelépést jelent.

Csepeli ma már a digitalizációt és az adatfigurációt egy **"techno-ökológiai környezet " létrehozásának tekinti.** A mesterséges intelligencia már nem csupán egy technológia, amely a társadalomban helyezkedik el. A társadalom maga egyre inkább egy mesterséges információs ökoszisztémában él.

Benne megjelenik az, amit Robot **sapiensnek nevez**.

Ezt a fogalmat gondosan meg kell különböztetni a Human 2.0-tól.

A robot sapiens a Homo sapiens mesterséges versenytársa: egy egyre autonómabb mesterséges ügynök, amely képes kognitívra és cselekvésre.

Az Ember 2.0 ehelyett az átalakult emberiség – a biológiai emberiség és a mesterséges intelligencia egyre növekvő integrációjának eredménye, amelyet potenciálisan a genetika, a nanotechnológia és a robotika kísér.

Így Csepeli jövője két különálló evolúciós lehetőséget tartalmaz:

verseny/helyettesítés: Homo sapiens kontra Robot sapiens;

és

Fúzió/átalakulás: A Homo sapiens Emberi 2.0-vá válik.

Ez a megkülönböztetés elméletének egyik legeredetibb szerkezeti jellemzője.

---

9. A "rendszer" és az "életvilág" átalakulása

A 2022-es tanulmány egy szociológiailag értékes megkülönböztetést is bevezet. Az MI egyszerre hat a **rendszerekre** és az **életvilágra**.

A rendszer szintjén átalakítja a gazdasági termelést, pénzügyet, adminisztrációt, politikai intézményeket, rendőrséget, katonai szervezést, közlekedést, egészségügyet és oktatást.

De az MI nem az intézményeknél áll meg. Áthatolja az élet világát – kommunikációt, barátságot, szexuális kapcsolatokat, szórakozást, fogyasztást, identitásformálást és mindennapi döntéshozatalt.

Ez létfontosságú, mert elméletileg egy társadalom formálisan demokratikus és gazdaságilag virágzó maradhat, miközben polgárai mindennapi valóságérzékelése egyre algoritmikusabb lett.

Az egyén technikailag megtarthatja a szabadságot, mikä:

egy ajánlórendszer határozza meg, mit lát;

egy platform határozza meg, kivel kommunikál;

egy navigációs rendszer határozza meg, hová megy;

egy minősítési rendszer határozza meg a hitelképességét;

egy MI tanácsadó alakítja döntéseit;

és a prediktív rendszerek határozzák meg, hogyan osztályozzák a szervezetek.

A formális autonómia tehát együtt létezhet a növekvő **funkcionális heteronomiával**.

---

10. Megfigyelés klasszikus diktatúra nélkül

Ez vezet Csepeli elméletének talán legerősebb politikai aspektusához.

A hagyományos totalitárius megfigyelésnek hatalmas gyengesége volt: drága volt. Az államoknak rendőrséget, informátorokat, bürokráciákat és fizikai archívumokat kellett alkalmazniuk.

A digitális megfigyelés ezt megfordítja.

Az emberek maguk hordozzák a megfigyelő eszközöket, előállítják a releváns adatokat, és gyakran önként engedélyezik azok gyűjtését ingyenes vagy kényelmes szolgáltatásokért cserébe.

Csepeli ezt a fejleményt Heidegger *Machenschaftjához* , majd később Shoshana Zuboff megfigyelés-kapitalizmus elemzéséhez köti. Az eredmény kialakuló rendszer átfogóbb lehet, mint a régi diktatúra, éppen azért, mert nem kell kényszerítőnek tűnnie.

A paradoxon így fejezhető ki:

a polgár együttműködik annak a szerkezet megalkotásában, amely átlátszóvá teszi.

Ez Csepeli "machináció" fogalmának olyan szociológiai specifikusságot ad, amely hiányzik egy tisztán heideggeri technológiakritikából.

---

11. MI és a politika átszervezése

Csepeli több szinten is elképzeli, hogy a mesterséges intelligencia több szinten lép be a politikai életbe.

Lehetővé teszi a választók részletes szegmentálását. Lehetővé teszi a személyre szabott politikai kommunikációt. Elősegíti a manipulációt szintetikus szövegek, képek és videók segítségével. Bővíti a kormányok képességét az állampolgárok besorolásában. És egyre inkább lehetővé teszi az adminisztratív rendszerek számára, hogy folyamatos emberi beavatkozás nélkül hozzanak döntéseket.

A 2022-es cikk már megvizsgálja annak lehetőségét, hogy a digitális adminisztráció hitelképesség, hűség, társadalmi viselkedés vagy egy tágabb "állampolgári érték" alapján értékelhetné az egyéneket. Az AI tehát potenciálisan átalakíthatja a politikai hatalmat az alkalmi beavatkozásból **folyamatos algoritmikus** értékeléssé.

Eközben a generatív MI fenyegeti magát a közszférát. Ha az autentikus és szintetikus beszéd, képek és videók egyre nehezebben megkülönböztethetők, a demokratikus vita episztemikus alapja instabilsá válik.

A kialakuló válság alapvetőbb, mint a szokásos félretájékoztatás. Ez a **társadalmilag intézményesített bizalom válsága a** valóságban.

---

12. A 2023-as ChatGPT-sokk: Csepeli álláspontjának jelentős radikalizálódása

Itt figyelhető meg valódi elméleti változás.

Egy 2023 áprilisi interjúban Csepeli kifejezetten kijelentette, hogy már nem lehet úgy gondolkodni az MI-ről, ahogy az ember csak egy-két évvel korábban tette. Elismerte, hogy korábbi feltételezése lényegében az volt, hogy az MI megkönnyíti vagy bonyolultabbá teszi az életet, de **az emberek szuverén** maradnak.

Ezután azt állította, hogy ez a feltételezés most tévesnek tűnik.

A GPT-típusú generatív modellek megjelenése Kurzweil szingularitását az ő véleménye szerint egy viszonylag távoli lehetőségről sokkal közelebb helyezte el.

Ez fontos bizonyíték arra, hogy Csepeli nem egyszerűen ismételte meg a 2020-as elméletét a ChatGPT után. **Átdolgozta a folyamat várható sebességét és mélységét**.

Számára a nyelvmodellek küszöböt jelentenek, mert korábban a nyelv volt kifejezetten az emberi kulturális intelligencia központi médiuma. Amint a gépek képesek koherens természetes nyelvű szövegeket alkotni, amelyek egyetemi szinten elfogadhatók, a társadalmilag hasznos intellektuális teljesítmény nagy része potenciálisan automatizálhatóvá válik.

Ezért előrejelzi a hagyományos szakmák gyors átírását vagy eltűnését, valamint az élethosszig tartó tanulás jelentős jelentőségét. Provokatív megfogalmazása, miszerint a jelenlegi tudás két éven belül használhatatlanná válhat, szociológiai figyelmeztetésként kell értelmezni a gyorsított elavulásra, nem pedig szó szerinti mennyiségi törvényként.

---

13. Egy társadalom, amelyben a legtöbb munka eltűnik

A 2023-as *Vissza a paradicsomba?* interjú jelentősen tovább viszi az érvet.

Csepeli szerint alig tud olyan irányt azonosítani, ahol az MI ne fejlődne; a digitális átalakulás egyre inkább "elragadja" gyakorlatilag az egész életet. Feltételezi, hogy azok a területek,

amelyek lerövidíthetetlenül függenek egy élő test fizikai jelenlététől – mint például a masszázs, balett, színház vagy testi intimitás – ellenállóbbnak bizonyulhatnak, miközben sok kézi és szellemi munka eltűnik a hétköznapi emberi foglalkozásból.

Ez az egyik legradikálisabb jövőbeli forgatókönyvét alkotja.

Az emberi élet két területre oszlik.

Az egyik egy rendkívül képességű **virtual-funkcionális gömb**, amelyet egyre inkább egy szuperagy vagy hálózatba kötött intelligencia irányít.

A másik az emberi **megtestesült szféra**, amely az érzés, a szexualitás, a szeretet, a közvetlen interperszonális jelenlét és élvezet maradéka.

Csepeli provokatívan hasonlítja az utóbbit a Paradicsomba való visszatéréssel: a kötelező munkától felszabadult emberek újra felfedezhetik a testesített és érzelmi létezést.

De ez szándékosan ambivalens volt. A "paradicsom" ugyanúgy lehet egy olyan függőség, amelyben az embereket anyagilag gondoskodják, de a civilizáció kognitív cselekvésének nagy részét a gépeknek adták át.

---

14. Az "Új Herceg": Csepeli legradikálisabb politikai MI-elmélete

A 2024-es *Új fejedelem* esszé véleményem szerint eddig Csepeli MI-elméletének csúcsformációja.

Kiindulópontja Nick Bostrom megkülönböztetése az MI **jóslatként**, **dzsinnként** és **szuverénként/uralkodóként működik** között. Csepeli azt állítja, hogy az adatficáció, digitalizáció, globális hálózatok és – egyértelműen – a 2022-től megjelenő generatív nyelvmodellek sokkal közelebb hozták a kollektív szuperintelligencia lehetőségét.

Az **ő "Új Hercege"** nem egy humanoid robot, és nem egyszerűen a ChatGPT nagyjából.

Ez egy **bolygói információs hatalomstruktúra** , amely akkor jön létre, amikor hatalmas adathalmazok, felhőalapú számítástechnika, hálózatok, generatív modellek és autonóm döntési rendszerek összekapcsolódnak.

Ez a kifejezés szándékosan Machiavelli kifejezését idézi.

A klasszikus herceg egy emberi uralkodó volt, aki területileg meghatározott államot irányított. Az új herceg a következők voltak:

nem biológiai;

hálózatba kötött;

potenciálisan globális;

folyamatosan tájékoztatva;

képes kombinálni az előrejelzést és az akciót;

és végül képes javítani saját működésének bizonyos részeit.

Egyesítené az orákulum tudásképességét, a dzsinn parancsok végrehajtását és a szuverén döntési képességét.

Ez Csepeli átmenete az **MI alkalmazások** elméletéből az **AI szuverenitás** elméletébe.

---

15. Tízéves távképe: nagyjából a 2030-as évek közepére

A cikk Csepeli egyik viszonylag ritka konkrét időhorizontját tartalmazza. Azt állítja, hogy körülbelül **tíz éven** belül alig marad olyan jelentős intézményi szféra, amelyet ez az "Új Herceg" érintetlenül érintett: a politika, a közigazgatás, a bíróságok, a rendőrség, az iskolák, kórházak, egyetemek, közlekedés és a hadsereg azok közé tartoznak, amelyeket említ. Mivel a szöveget 2023-ban készítették és 2024-ben adták ki, ez nagyjából **2033–2034-re** egyezik.

De ezt nem szabad félreérteni, mert "Csepeli szuperintelligencia jósolja 2034-re."

Nem állapít meg pontos AGI vagy szingularitási dátumot.

Írásaiban ehelyett három különböző időbeli állítás található.

2020-ban **Kurzweil 2045-ös szingularitását** tárgyalta, kifejezetten lehetővé téve a korábbi vagy későbbi felismerést.

A ChatGPT után, 2023-ban lényegesen közelebbnek ítélte a szingularitást.

2024-ben körülbelül egy évtizedet jósolt előre a fejlett MI szinte univerzális **intézményi** penetrációjának.

Ezek különböző állítások.

Az utóbbi jóval kevésbé szélsőséges, mint azt állítani, hogy 2034-re szükségszerűen létezni fog egy teljesen autonóm szuperintelligencia.

---

16. MI vállalatok mint új geopolitikai szereplők

Az Új Herceg *másik fontos kiterjesztése* a szuverenitásról szól.

A határterületi MI infrastruktúrája – hatalmas adatok, adatközpontok, fejlett chipek, felhőalapú számítástechnika, modellképzés és globális hálózatok – korlátozott számú vállalatban és államban koncentrálódik.

Ennek következtében Csepeli azt állítja, hogy a nagy technológiai vállalatok valamiféle **geopolitikai cselekvőképességet** szereznek. A nemzetállamok formálisan is szuverén maradhatnak, miközben technológiailag olyan magáninfrastruktúráktól függnek, amelyeket maguk nem tudnak reprodukálni.

Csepeli meglehetősen messzire viszi a geopolitikai érvet, az Egyesült Államokat és Kínát a kialakuló MI rend fő lehetséges pólusaiként mutatja be.

Ez az elmélet része koncentrációs szinten hihető, de legerősebb bináris formájában még megkérdőjelezhetőbb. A jövőbeli MI hatalmi struktúrája magában foglalhat transznacionális vállalati hálózatokat, több államot, nyílt forráskódú ökoszisztémákat és változó félvezető-, energia- és adatszövetségeket is. Egy egyszerű amerikai–kínai zéró összegű modell veszélyezteti az intézményi összetettség alábecsülését.

Ennek ellenére nagyobb szociológiai pontja fontos: a **fejlett MI megváltoztatja a szuverenitás jelentését, mert maga a számítási kapacitás stratégiai erőforrássá válik, amely összehasonlítható az ipari, katonai és pénzügyi hatalommal.**

---

17. Ember 1.0, Robot sapiens, Ember 2.0 és az Új Herceg

Csepeli érett elmélete tehát négy ideális típussal tisztázható.

Koncepció	Jelentés
Human 1.0	A biológiai-kulturális Homo sapiens történelmileg a mesterséges intelligenciával való mély integráció előtt alakult ki.
Robot sapiens	Autonóm mesterséges intelligencia, amely megtestesítette vagy más módon képes kognícióra és cselekvésre, és az emberek funkcionális versenytársává válik.
Human 2.0	Egy átalakult ember, amelyet a biológiai intelligencia egyre intimebb integrációja hozott létre az MI-vel, és végső soron a genetikai mérnökséggel, a nanotechnológiával és a robotikával.
Új Herceg	Kollektív, hálózatos, potenciálisan bolygói szuperintelligencia, amely infrastrukturális, és talán végül szuverén hatalmat gyakorol a társadalmi rendszerek felett.

Ez a négyrészes különbség teszi Csepeli beszámolóját jóval kifinomultabbá, mint egy egyszerű "robotok fognak helyettesíteni minket" narratívát.

A robot sapiens a helyettesítést jelenti.

A Human 2.0 a hibridizációt jelképezi.

Az Új Herceg a rendszerszintű koncentrációt és szuverenitást jelképezi.

És a **Human 1.0** az a történelmi forma, amelyet mindhárom kihívás elé állít.

---

18. Transhumanizmus: Az AI végül átalakítja magát az emberi testet

A 2025-ös *A transzhumán művészet tragédiája* című esszé további változást jelez.

Csepeli egyre inkább az MI-t egy nagyobb konvergens technológiák csomagjába ágyazza: **mesterséges intelligencia + genetikai manipuláció + nanotechnológia + robotika + virtuális** környezetek.

A transzhumanizmus definíciója szerint technológia révén kiterjeszti az emberi létezés természetes és társadalmi határait. A hálózatba kapcsolt információk egyre inkább destabilizálják azt, ami az emberi identitás végső fix pontjának tűnt: az élő testet. A nanotechnológia, a genetika és a robotika még a testet is technológiailag módosíthatóvá teheti, míg a virtuális környezetek gyengítik a fizikai és pszichológiai jelenlét régi különbségét.

Ez egy jelentős elméleti bővítés.

A korábbi MI probléma a következő volt:

Mi történik, ha a gépek intelligenssé válnak?

A későbbi Csepeli kérdés így alakul ki:

Mi történik, amikor a gép és az ember közötti különbség megszűnik stabil lenni?

Ez egy igazán transzhumán, nem pusztán technológiai kérdés.

---

19. MI és művészet: az emberi egyediség visszahúzódó határa

Különösen érdekes fejlemény történt a 2022 és 2025-ös szakaszok között.

Korábban Csepeli hajlamos a művészi kreativitást, a megtestesült jelenlétet, az érzelmeket és az értéket az emberi egyediség menedékeiként kezelni.

A generatív mesterséges intelligencia még ezen a határon is nyomást gyakorol.

2025-re elfogadja, hogy a mesterséges intelligencia magát a művészeti termelést is áthatolja. A szöveg, képek, zene és más kulturális termékek egyre inkább generatív rendszerek révén készülhetnek. Ennek következtében az emberi védelmi vonal egyre mélyebbre kerül.

A döntő megkülönböztetés már nem egyszerű:

Gépi számítás kontra emberi kreativitás.

Ez így alakul:

Szintetikus kulturális termelés a megtestesített emberi létezéssel, megélt tapasztalattal, igazsággal, jelentéssel és egzisztenciális jelenléttel szemben.

Heideggeri kifejezéssel Csepeli attól tart, hogy a technológiailag generált kreativitás technikailag zseniálissá válhat, miközben elveszíti a művészet, az igazság és az emberi létezés közötti kapcsolatot.

Ez a változás valami fontosat mutat a gondolkodásának fejlődésében: **amikor a technológiai fejlődés aláásja egy feltételezett kizárólagos emberi képességet, Csepeli inkább az emberi magot helyezi át, ahelyett, hogy tagadná a technológiai fejlődést.**

---

20. Mi az igazán eredeti Csepeli elméletében?

Pontatlan lenne ezeket a fogalmakat Csepeli találmányainak nevezni.

A szingularitás főként Kurzweilből származik.

A szuperintelligencia és az orákulum/dzsinn/szuverén taxonómia Bostromból származik.

A megfigyelő kapitalizmus erősen köszönheti Zuboffnak.

A kalkulatív uralom technológiai kritikája Heideggertől ered.

Az emberfeletti horizont Nietzschét idézi.

A hálózati osztályú érveknek elődei vannak a netokrácia elméleteiben.

Sokkal jellegzetesebb, az **Csepeli szintézise**.

Öt különösen fontos hozzájárulást neveznék ki saját intellektuális felépítéséből.

Először az **adatfikációt → digitalizációt → a hálózatépítést, →→ a mesterséges intelligenciát társadalmi hatalmat** egyesíti egy egységes szociológiai sorrendben.

Másodszor, az **episztemológiai/axiológiai megkülönböztetés** eredeti antropológiai központot ad az AI elméletének: a gépek fokozatosan magukhoz foglalhatják a tudást, miközben a jelentés kérdését megoldatlanul hagyják.

Harmadszor, **összekapcsolja a Robot Sapiens-t, az Ember 2.0-t és az Új Herceget** , mint három lehetséges MI fejlődési következményet: mesterséges versenytársat, átalakult emberi intelligenciát és a kollektív szuverén intelligenciát.

Negyedszer, Heideggeri **gépészeti koncepciója az önkéntes digitális megfigyeléssel kombinálva** erőteljes magyarázatot ad arra, hogy miért nem kell a jövőbeli teljes irányítási rendszernek hasonlítania a huszadik századi diktatúrára.

Ötödszer, az MI által okozott munkanélküliséget egy mélyebb kérdéshez köti, mint a jövedelemelosztás: **mi lesz az emberi jelentéssel, amikor a gazdaságilag szükséges munka megszűnik az élet szervezésében?**

Ezek azok a szempontok, amelyekben Csepeli írásai a leginkább túlmutatnak egy nemzetközi MI-jövős irodalom gyűjteményén.

---

21. A legmélyebb ellentmondás: episztemikusan felsőbbrendű, de axiológiailag üres MI

Talán a legeredményesebb módja Csepeli egész projektjének rekonstrukciójának egyetlen paradoxon köré épül.

Egy jövőbeli mesterséges intelligencia több ténytudással rendelkezhet, mint bármely ember, jobban diagnosztizálhatja a betegségeket, optimalizálhatja a termelést, hatékonyabban írhat, jobban irányíthatja a szervezeteket, jobban előrejelzi a viselkedést, és talán tudományos kutatásokat is végezhet.

Mindezek logikusan nem ad választ a kérdésre:

Mire kellene ezt az intelligenciát felhasználni?

Csepeli számára a technológiai civilizáció **veszélyezteti az episztemológiai mindenhatóságot az axiológiai ürességgel kombinálva**.

Ez magyarázza Kant-elméleti ismételt nietzschei újrafogalmazását.

A felvilágosodás imperatiiva a **"Sapere aude" volt — merd tudni**.

Amint a gépek rendkívül jóvá válnak a tudásban, a specifikus emberi imperatív egyre inkább olyasmi lesz, mint **"merj élni".**

Az emberi feladat az információgyűjtéstől az érték, ítélőképesség, megkérdőjelezés, egzisztenciális elkötelezettség és ellenállás fenntartásáig változik. A 2022-es tanulmány kifejezetten bemutatja ezt a teljes gépészeti rendezéssel szembeni ellenállást, mint a technológiailag lehetővé tett megfigyelés ellensúlyát.

---

22. Felszabadítás kontra dominancia

Ezért Csepelinek nincs egyszerű disztópikus MI-elmélete.

Két jövő nyitva marad munkáinak nagy részében.

Az **emancipációs pálya** felszabadulást kínál az egyhangú munkától, hatalmas javulást az egészségügyben, hosszabb élettartamot, biztonságosabb termelést, nagyobb bőséget, egyéni szolgáltatásokat, drámaian bővített tudást, és talán több időt a családra, barátságra, szeretetre, kreativitásra és elmélkedésre.

A **dominancia teljes** viselkedési átláthatósághoz, algoritmikusan strukturált választásokhoz, hatalmas technológiai egyenlőtlenséghez, platformtulajdonosoktól való függőséghez, az emberi kognitív autonómia gyengítéséhez, és végül egy olyan kollektív intelligenciához vezet, amelynek összetettsége miatt az értelmes emberi irányítás lehetetlenné válik.

Ugyanaz a technológia mindkét forgatókönyvet támogatja.

Ezért nem az, hogy "Az MI jó vagy rossz?"

Ez a következő:

Milyen társadalmi kapcsolatok, tulajdonjogi struktúrák, politikai intézmények és emberi életformák veszik majd körül az MI-t?

Ez alapvetően szociológiai kérdés.

---

23. Ahol Csepeli előrejelzései a legerősebbek

Az elmélet több összetevője különösen jól öregedett.

Korai ragaszkodása ahhoz, hogy **az adatok koncentrációja, nem csak az algoritmus, az MI hatalom alapja**, most fontosnak tűnik.

Az az elvárása, hogy az automatizálás a kézi munkáról a kognitív munkára terjed, egyértelműen előre jelezte a 2022 utáni generatív MI problémát.

Figyelmeztetése, miszerint a mesterséges szöveg és képek megingathatják a közszféra bizalmát, sokkal konkrétabbá vált.

A technológiai vállalatokra, mint kvázi-politikai szereplőkre való fókusza előre jelezte a felhőinfrastruktúra, az alapmodellek és a számítási kapacitás növekvő stratégiai jelentőségét.

És a technológiai munkanélküliség mint jövedelemprobléma és a munkanélküliség mint **jelentésprobléma közötti** megkülönböztetése szociológiailag mélyebb, mint a legtöbb hagyományos munkaerőpiaci előrejelzés.

---

24. Ahol elmélete spekulatívabb

A radikális forgatókönyveket azonban el kell választani a szilárdabb társadalmi elemzéstől.

Csepeli néha meglehetősen gyorsan átlép a mai gépi tanulástól az "erős MI", a robot sapiens, a szingularitás, az önfejlesztés és a szuperintelligenciához. Ezek a szakaszok technológiailag nem azonosak, és az egyik sem jelenti automatikusan a következőt.

Hasonlóképpen, ha az MI, a génszerkesztés, a robotika és a nanotechnológia egyetlen emberi 2.0 fogalmába egyesítik, az meggyőző evolúciós narratívát ad, de egyetlen konkrét technikai mechanizmus sem mutatja ki, hogy ezeknek a különálló technológiáknak egy integrált poszthumán organizmussá kell össze kell állniuk.

A hatalmas munkaerő-elmozdítás várakozása forgatókönyvként hihető, de a megsemmisült, átalakult foglalkozások és újonnan létrehozott munkahelyek pontos egyensúlya empirikus kérdés.

A gépek Turing-teszten való átmenéséről vagy az autonóm intelligencia megközelítéséről szóló nyelvezete részben filozófia: az emberszerű nyelv sikeres létrehozása önmagában nem teremti meg a tudatosságot, a megértést, a szándékosságot vagy az általános intelligenciát.

És az "Új Herceg" leginkább **ideálszerű figyelmeztetésként értelmezhető az algoritmikus hatalom koncentrációjáról**, nem pedig egy bizonyított előrejelzésként, hogy egy szó szerint egységes bolygói MI-szuverénnek kell megjelennie.

---

25. Mennyire radikális Csepeli előrejelzése a következő évekre?

Gondolkodásában észrevehető felgyorsulás van:

2015: a társadalom egyre inkább adatosan.

2018: a hálózatba épült MI fenyegeti az emberi episztemológiai munka nagy részét, és a szingularitás kezd megjelenni a horizonton.

2020: Az Emberi 2.0 és a szuperintelligencia rendszerszerű, távoli jövőbeli lehetőséggé válik.

2022: A robot sapiens mind a társadalmi rendszereket, mind az életvilágot képes áthatolni.

2023: a generatív MI kifejezetten arra készteti, hogy újragondolja az emberi szuverenitás fenntartásának feltételezését.

2024: a kollektív szuperintelligencia válik az "Új Hercegnek", és közel egy évtizeden belül előrejelzett az intézményi MI szinte univerzális behatozata várható.

2025: még a test, a kreativitás és a művészet is beépül egy transzhumán átalakulásba.

Így Csepeli saját elméletének pályája is jelentős. **A várható átmenet egyre közelebb kerül, és a kizárólag emberi maradt szféra egyre szűkebb lesz.**

Az MI-vel kapcsolatos munkák közül, amelyeket 2026 szeptemberéig ellenőrizhettem, a 2025-ös transzhumanizmus esszé ennek az intellektuális pályának a legfrissebb jelentős kiterjesztése; a 2024-es *New Prince* továbbra is a legvilágosabb, kompakt megfogalmazása a közelgő politikai-társadalmi rendnek.

---

26. Általános értékelés

Csepeli kidolgozta az egyik legátfogóbb magyar szociálelméleti értelmezést a mesterséges intelligenciáról. A döntő eredmény nem egy új műszaki elmélet az MI-ről, hanem az MI új **civilizációs környezetként való** megértésének kísérlete.

Elmélete a következő oksági architektúrára redukálható:

az adatfigurázás és a digitalizáció biztosítja a nyersanyagot;

a hálózatok és a felhőinfrastruktúrák biztosítanak bolygói kapcsolódást;

a tanulás és a generatív MI egyre autonómabb kognitív teljesítményt nyújt;

Ezek a technológiák áthatolnak **a gazdaságban, a munkában, a közigazgatásban, a politikában, a médiában, az oktatásban, az egészségügyben, a biztonságban és a mindennapi** életben;

az infrastruktúra feletti ellenőrzés új **osztály- és politikai hatalomelosztást teremt**;

a sok szükséges munka eltűnése veszélyezteti a munká**, a jövedelem, a státusz és az jelentés hagyományos** kapcsolatát;

az átfogó adatfeldolgozás lehetővé teszi az **önkéntes teljes megfigyelést**;

a fejlett MI vagy **Robot Sapiens**, **Human 2.0**, vagy egy hálózatba kötött **Új Herceg felé** fejlődik;

és végső soron az emberiség szembesül a géppel hozzáférhető **episztemológiai intelligencia** és az axiológiai jelentés még emberi problémája között.

Ez az utolsó elem adja Csepeli elméletének jellegzetes jellegét. Számára a végső veszély nem az, hogy a gépek feltétlenül gonoszsá váljanak. Az, hogy egy civilizáció meglepően intelligenssé válhat instrumentális értelemben, miközben elveszíti a képességét megkérdezni **, miért** akarja azt, amit a gépei rendkívül képesek elérni.

Ezért Csepeli elméletét leginkább nem hagyományos technológiai determinizmusnak, sem hagyományos MI-kockázatelméletnek lehet besorolni. Ez egy **evolúciós, szociológiai és egzisztenciális elmélet a digitális társadalomból egy potenciálisan transzhumán civilizáció felé való átmenetről**. Empirikus előrejelzései egyenetlenek maradnak, és néhány szuperintelligencia forgatókönyve szükségszerűen spekulatív, de mint kísérlet arra, hogy az MI-t egyszerre kapcsolja össze társadalmi rétegződéssel, politikai hatalommal, megfigyeléssel, munkával, kultúrával, emberi identitással és az emberi létezés lényegével, rendkívül átfogó.

A legjelentősebb fejlemény a *Human 2.0* után tehát nem csupán az, hogy Csepeli "pesszimistábbá" vált. Inkább **a generatív MI után rövidítette a várható távolságot a mai digitális társadalom és az emberi kognitív szuverenitás megkérdőjelezhető küszöbje** között. A 2024-es *New Prince* a radikalizáció legvilágosabb kifejezése, míg a 2025-ös transzhumanizmus esszéje a társadalomból az emberi lény alkotmányába terjed.

****

II. Lengyelország

****

Andrzej Zybertowicz elmélete

Kérjük, részletesen írja le és elemezze Andrzej Zybertowicz tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint terveztek.

Andrzej Zybertowicz az egyik közép-európai szociológus, aki megpróbálta a mesterséges intelligenciát nemcsak új technológiaként, sem elsősorban foglalkoztatási, magánéleti vagy szabályozási problémáként kezelni, hanem az **emberi civilizáció alapvető architektúrájának lehetséges** átalakításaként. Hivatalos életrajzi profilja most már kifejezetten azonosítja az új technológiák, különösen az MI társadalmi következményeit, mint legújabb kutatásainak egyik fő témáját.

Van azonban egy fontos módszertani minősítés. Zybertowicz **nem alkotott formális előrejelző elméletet az MI-ről, amely összehasonlítható lenne a technológiai fejlődés matematikai modelljével**, és nem vezetett egy nagy empirikus kutatási programot, amely az MI hatásait adatbázisokon keresztül mérte. Munkáját inkább úgy írható le, mint **a civilizációs kockázat, hatalom és technológiai átalakulás makroszociológiai elmélete**. Valóban, az *AI Eksploracja* című művében ő és Katarzyna Zybertowicz kifejezetten elismerik, hogy témájuk meghaladja a hagyományos szociológiai kompetenciát, sok alproblémát csak részben lehet feltárni, ezért a **"civilizációs biztonságot"** tekintik átfogó nézőpontként.

Értékelésem szerint elmélete egy koherens sorozatként rekonstruálható, amely nagyjából egy évtizedre nyúlik fel:

Korszak/munka	Fő elméleti lépés
2015 – *Samobójstwo Oświecenia?*	A technológiai fejlődési paradoxon: az emberi racionalitás olyan rendszereket hozhat létre, amelyek elpusztítják az emberiség képességét, hogy megértse és irányítsa saját világát.
2022 – *Cyber kontra real*	A digitális kapitalizmus, a platformok, a megfigyelés, a polarizáció és a kognitív degradáció teremti meg azt a társadalmi környezetet, amelybe a fejlett MI belép.
2023–2024 – "Két görbe" szakdolgozat	Az MI kognitív képességei nőnek, miközben az emberek kollektív kognitív képességeit állítólag gyengíti a digitális környezet.

Korszak/munka	Fő elméleti lépés
2024 – *AI Eksploracja*	Teljes civilizációs kockázatelmélet: MI mint erősokszorzó, emergent fekete doboz, geopolitikai fegyver, radikális aszimmetriák teremtője és potenciálisan egzisztenciális fenyegetés.
2025 – *Trwała mapa dla niestabilnego bezpieczeństwa*	Az MI elméletet beépítve a nemzetbiztonsági és geopolitikai elemzésbe, beleértve a 2–5 és tízéves horizontokat is.
2026 – technológiai munkanélküliségi elemzések	Nagyobb figyelem az emberi munkaerő gyors helyettesítésére MI modellekkel és robotikával.

Különösen érdekes, hogy Zybertowicz elméletének alapvető felépítése már **a ChatGPT előtt is jelen** volt. A generatív MI forradalom nem arra késztette őt, hogy új világnézetet találjon ki; inkább váratlanul erős empirikus indokot adott egy olyan diagnózisára, amelyet legalább 2015 óta dolgozott ki.

---

1. A kiindulópont: *Samobójstwo Oświecenia?* (2015)

Az alapvető munka a kollektív kötet

Andrzej Zybertowicz Maciej Gurtowski, Katarzyna Tamborska, Mateusz Trawiński és Jan Waszewski társaságában, *Samobójstwo Oświecenia? Jak neuronauka i nowe technologie pustoszą ludzki świat* [*A felvilágosodás öngyilkossága? Hogyan pusztítják el az idegtudomány és az új technológiák az emberi* világot], 2015.

Ez egy 522 oldalas kollektív tanulmány volt, amelyet Zybertowicz vezetett. Zybertowicz maga írta a bevezetőt, valamint más részeket a felvilágosodásról, automatizálásról és digitális forradalomról, valamint a technotudományról és a hatalom jövőjéről szóló fejezeteket. A kötet foglalkozott idegtudománysal, megfigyeléssel, Big Data-val, agy-számítógép interfészekkel és transzhumanista technológiákkal is.

Ez azért fontos, mert Zybertowicz későbbi MI-elmélete nem érthető meg a korábbi keretrendszer nélkül.

1.1 A "felvilágosodás öngyilkossága" tézise

Kiindulópontja egy paradoxon.

A felvilágosodás azt ígérte, hogy az emberi ész megszabadulása a hagyománytól, a babonáktól, vallástól és az örökölt tekintélytől lehetővé teszi a tudomány számára, hogy felfedezze a világ igazságát, majd a racionális emberek szándékosan egyre virágzóbb, igazságosabb és irányíthatóbb társadalmakat építhetnek.

Zybertowicz azt állítja, hogy a technológiai modernitás pont az ellenkező eredményt hozhatja.

A tudományos ész intézményesen kapcsolódik hozzá:

a tudomány → technológiai innováció → piacok → vállalatok → verseny → folyamatos innováció.

Így az innováció már nem elsősorban a következő kérdés határozza meg:

Milyen emberi világot kellene építenünk?

és egyre gyakrabban:

Mit lehet feltalálni, kereskedelmi forgalomba hozni, telepíteni vagy militarizálni, mielőtt a versenytársak ezt megtennék?

Megfogalmazása szerint a felvilágosodás kultúrája által felemelt felszabadult ész így a piaci erők, a kapitalizmus és a tömegfogyasztás szolgájává válhat; új újításokat vezetnek be, mielőtt a társadalom megértené a korábbi következményeket.

Ez a Zybertowicz mesterséges intelligencia szociológiájának első kulcsfontosságú mechanizmusa:

A technológiai fejlődés sokkal gyorsabb, mint a társadalom reflektív önkontroll képessége.

Nem az egyéni technológiai találmányok jelentik a legmélyebb veszélyt. A **társadalmi mechanizmus hozza létre az innovációkat, mint ahogy az intézmények, a kultúra és az emberi kogníció képes lenne őket magába szívni**.

---

2. MI és az "új sötét korok" lehetősége

A 2015-ös könyv legfigyelemreméltóbb része, hogy Zybertowicz már nyíltan tárgyalt valamit, ami hasonlít a mai AGI/szuperintelligencia problémára.

Azt állítja, hogy ha egy mesterséges általános intelligencia képes lenne önállóan tanulni és fejlődni, akkor elkezdhetne matematikai vagy tudományos reprezentációkat építeni az emberi kognitív képességen túl. Az emberi tudósok így nem tudnák megérteni sem az AI által előidézett felfedezéseket, sem azokat a belső elveket, amelyek szerint maga a rendszer fejlődik.

Ez vezet ahhoz, ami valószínűleg a 2015-ös elmélet legfontosabb és legeredetibb metaforája:

A technológiai "új sötét korszakok".

Történelmileg a tudományos fejlődés növelte az emberi érthetőséget:

ismeretlen világ→ kutatás → tudományos magyarázat → nagyobb emberi megértés.

A fejlett MI visszafordíthatta a nyilat:

az emberi tudás → MI által létrehozott tudás → emberfeletti összetettség → csökkenő emberi érthetőség.

Ha egyre autonómabb MI-rendszerek lennének felelősek a tömegesen digitalizált társadalom koordinálásáért, a civilizáció így rendkívül hatékonysá válhatna, miközben **alapvetően érthetetlenné válna az benne élő emberek** számára.

Zybertowicz élesen megfogalmazza a lényeget: évszázadok tudományos kísérleteinek eredménye a valóság megvilágítására egy olyan világ lehet, amelyet olyan mechanizmusok irányítanak, amelyeket elvben egyetlen emberi elme sem tudna felfogni.

Ez az elképzelés jóval kifinomultabb, mint az egyszerű állítás, hogy "az MI erősebb lehet az embernél."

A probléma nem csupán **a hatalmi fölény**.

Ez az **epitemikus aszimmetria**:

$$\text{complexity of the system} > \text{human capacity to comprehend the system.}$$

Ha ezt a küszöböt átlépik, a demokrácia, az elszámoltathatóság és a politikai kontroll mély problémával szembesül: hogyan tudnak az emberek értelmesen kormányozni egy olyan rendszert, amelynek oksági működését nem értik?

---

3. Technoscience az autonóm "technológia" helyett

Egy másik központi jellemző megkülönbözteti Zybertowiczt a technológiai deterministáktól.

Nem mondja egyszerűen:

A technológia önállóan fejleszti és átalakítja a társadalmat.

Ehelyett egy technotudományi szociológiát javasol.

A technológiai fejlődés a tudomány, vállalatok, piacok, államok, katonai verseny, kutatószervezetek, befektetők és több ezer decentralizált innovátor kölcsönhatásából ered. A piac rendkívül erős, de a gazdag államok és vállalatok is részt vesznek abban, amit ő egy új kognitív vagy technológiai fegyverkezési versenynek nevez.

Így nincs egyetlen "kapitány", aki irányítaná a technológiai civilizációt.

Ez egy kontrollparadoxont eredményez:

sok szereplő irányítja az egyéni technológiai projekteket, míg senki sem irányítja a technológiai civilizáció összesített irányát.

A Google irányíthat egy rendszert; egy kormány szabályozhatja a másikat; egy katonai laboratórium irányíthatja a programját; a befektetők megszüntethetik vagy bővíthetik az egyes projekteket.

Mégsem olyan globális szereplő, aki képes lenne feltenni:

Folytatódjon-e a teljes technológiai folyamat ebben az irányban?

Ez kulcsfontosságúvá válik, amikor az MI geopolitikai versenybe lép.

---

4. A korai elmélet már integrálta az MI-t az idegtudománysal, a Big Data-val és a megfigyeléssel

A 2015-ös könyv másik jellemzője az volt, hogy ragaszkodott hozzá, hogy a technológiákat nem szabad külön kezelni.

Az architektúrája összekapcsolja az idegtudományt, MI-t, Big Data-t, automatizált megfigyelést, neurotechnológiát, agy-számítógép interfészeket, automatizálást és vállalati hatalmat. A tartalomjegyzék maga mutatja a tervezett szintézist: a fejezetek az idegtudománytól és az emberi természettől az automatizáláson, megfigyelésen és transzhumanista katonai technológián át a "Technoscience és a hatalom jövője" felé haladnak.

Zybertowicz ezért előre látta a témát, amely később sokkal fontosabbá válik későbbi munkáiban:

A legnagyobb civilizációs hatások nem az AI elszigetelődéséből fognak eredni, hanem **több technológia közötti szinergiákból**.

Az MI + idegtudomány + mindenütt jelenlévő okostelefonok + Big Data + közösségi média + megfigyelés külön-külön minőségileg eltérőt hozhat elő.

Ez 2025-re válik a **kognitív hadviselés** elméletévé.

---

5. *Cyber kontra real* (2022): a társadalom felépítése, amelybe az MI érkezik

A következő nagy szakasz a *Cyber kontra real. Cywilizacja w techno-pułapce* [*Cyber versus Real: Civilizáció a Techno-Trapban*], egy 2022-es könyvhosszú beszélgetés Zybertowicz és Jarema Piekutowski szociológus között.

Az építészete feltáró. A szekciók címe*: Techno-forradalom*, *Techno-munka*, *Techno-ember*, *Techno-polarizáció*, *Techno-állapot*, *Techno-világ*, *Techno-biztonság*, *Techno-háború*, *Techno-remény, vagy a Nagy Lassulás*, végül pedig *Valós kontra Cyber*.

Ezért az MI egy sokkal nagyobb digitális civilizációs elméletbe épül.

A legfontosabb lépés 2015-től 2022-be a következő:

2015-ben Zybertowicz elsősorban a következőket kérdezi:

Mi történik, ha a technológiai fejlődés túllépi az emberi irányítást?

2022-re egyre inkább felteszi a kérdést:

Milyen embereket és társadalmakat hoz létre a digitális forradalom, mielőtt még a nagyon fejlett MI megérkezne?

Ez a változtatás döntő jelentőségű.

---

6. A "szerkezetileg mérgezett infoszféra"

A központi köztes változó az információs környezet.

Zybertowicz végül a kortárs közkommunikációt " **szerkezetileg mérgezett infoszférának" nevezi.**

Az AI Eksploracja esetében a fogalom viszonylag pontos definíciót kap. Ez egy olyan közkommunikációs környezetre utal, amelyben az emberek egyre nehezebben tudják megkülönböztetni:

Igazság a hamisságtól,
fontos, a lényegtelen,
sürgős a nem sürgőstől,
barát az ellenségtől.

Ezt az állapotot az információtúlterhelésnek, a folyamatos széttagolt információáramlásnak, a félretájékoztatásnak, függőséget okozó digitális mechanizmusoknak, a sebességnek és a közösségi média rendszerek által generált multitaskingnek tulajdonítja.

Ez a fogalom kritikus, mert Zybertowicz **nem** képzeli el, hogy az MI belépne egy egyébként egészséges liberális-demokratikus információs rendbe.

Oksági sorrendje nagyjából:

digital platforms → attention economy → information overload → cognitive fragmentation → polarization → weakened social trust → reduced collective rationality.

Csak ekkor érkezik meg egyre erősebb MI.

Ezért az elmélete sokkal sötétebb, mint azok, amelyek 2022-ben ChatGPT-vel kezdik elemzésüket.

---

7. Techno-lelkesedés mint ideológia

Egy másik fontos fogalom, amelyet a *Cyber kontra real* és az *AI Eksploracja radikalizált* , a **Techno-Entuzjazm** – techno-lelkesedés.

Zybertowicz nem csupán a technológia kedvelésének határozza meg, hanem mint egy szokás vagy ideológiai hajlamot, amely a egymást követő technológiákat a fejlődés megnyilvánulásaiként fogadja el anélkül, hogy megfelelően vizsgálná a rendszerszintű következményeket. *A Cyber kontra real című* művében ezt a technológiai fejlődésbe vetett kritikátlan bizalomként írja le, amely aláássa a kritikus reflexió képességét.

Az *AI Eksploracja* egy egész fejezet címe:

"Techno-Entuzjazm – a legveszélyesebb kortárs ideológia."

Miért "ideológia"?

Mert Zybertowicz számára ideológiai funkciót lát el, amely analóg a régebbi legitimáló hitrendszerekkel: egy történelmileg feltételes elrendezést természetesnek és elkerülhetetlennek lát.

A tipikus techno-lelkes javaslat a következő:

A technológia úgyis fejlődni fog; ezért a társadalomnak alkalmazkodnia kell.

Zybertowicz ezt megfordítja:

Ha az emberek létrehoztak technológiai intézményeket, joguk van eldönteni, mely technológiai pályákat engedélyezik.

Ez az egyik oka annak, hogy a javaslata végül **Lassú** MI-vé vagy technológiai moratóriummá válik.

---

8. A "két civilizációs görbe": valószínűleg legemlékezetesebb MI tézise

2023 körül Zybertowicz két **görbe metaforáján keresztül kezdte kifejezni** elméletét.

Egy görbe emelkedik:

$$AI\ cognitive\ capability \uparrow$$

A másik elutasítás:

$$human\ collective\ cognitive\ capability \downarrow$$

Azt állította, hogy a digitális média csökkenti a tartós koncentráció és a "mély munka" kapacitását, miközben az MI rendszerek egyre összetettebb intellektuális feladatokban gyorsan fejlődnek. Ezért figyelmeztetett arra, hogy a két görbe metszéspontján találkozhat.

Ez érdekesebb, mint a hagyományos "MI okosabb lesz az embereknél" tézis, mert ez teszi a küszöbértéket **relációsá**.

Az embernek nem kell mindentudó gépet létrehozniuk.

Tegyük fel, hogy az MI képességei 40-ről 70-re lépnek, miközben az emberi kollektív kognitív kompetencia 100-as marad. Az emberiség továbbra is domináns.

De ha az MI 40-ről 70-re emelkedik, miközben az emberi összetett rendszerek megértési képessége 100-ról 65-re csökken, a relatív erőviszony sokkal korábban megfordulhat.

Szimbolikusan:

$$A(t) \uparrow,\ H(t) \downarrow$$

Az irányítás elvesztése akkor fordulhat elő, amikor

$$A(t) \geq H(t),$$

Még ha nem is érte el semmit, ami távolról is hasonlítana a végtelen vagy isteni intelligenciára.$A(t)$

Ez Zybertowicz egyik legerősebb érvét eredményezi:

Az MI-nek nem kell tudatosnak, teljesen autonómának vagy univerzálisan szuperintelligenssé válnia ahhoz, hogy rendkívül veszélyessé váljon.

Egy rendszernek csak annyit kell teljesítenie, hogy felülmúlja az embereket egy stratégiailag kritikus területen – pénzügy, kiberhadviselés, katonai célzás, meggyőzés, megfigyelés, tudományos kutatás vagy politikai manipuláció. 2023-as interjújában kifejezetten azt állította, hogy egy MI-nek nem kell tudatosnak, önmagát tanulónak vagy teljesen szuperintelligensnek lennie ahhoz, hogy "szuperveszélyessé" váljon; egy kritikus fontosságú területen a felsőbbrendűség is elegendő.

Ez jelentős eltérés a népszerű vitáktól, amelyek szinte kizárólag a tudatos AGI-re összpontosítanak.

---

9. *AI Eksploracja* (2024): az érett rendszer

A legszisztematikusabb bemutató az *AI Eksploracja. Andrzej Zybertowicz odpowiada Katarzynie Zybertowicz*, amelyet először 2024-ben adtak ki. A Lengyel Nemzeti Könyvtár egyszerre sorolja be a civilizáció, informatizáció, információs társadalom, mesterséges intelligencia, digitális technológiák és társadalmi változások kategóriákba.

A szerkezet láthatóvá teszi az elméleti fejlődést.

Az első rész, **a "Körbenézés"** azt kérdezi, milyen civilizációba lép be az MI.

A második rész, **a "Hamis megnyugtató elméletek"** bemutatja az erőszorzót, a megjelenést és az állítólagos hamis analógiát, miszerint "az emberiség mindig is megbirkózott az új technológiákkal."

A III. rész, **a "We See It Darkly"** a techno-lelkesedést, az MI fegyverkezési versenyt és a negatív forgatókönyveket vizsgálja.

A IV. rész, **a "Remény részei" a** "Időt nyerjetek!" **javaslatában csúcsosodik ki.**

Az epilógus ezután a rendszerarchitektúrára terelődik: globális technológiai kudarcok, az internet mint az emberiség Achilles-sarka, technológiák szétválasztása, miközben az embereket köti össze, és valami olyasmit építenek, ami a technológiai területek közötti "határátlépésekhez" hasonlít.

Ez már nem pusztán kulturális kritika. Ez egy kezdeti **elmélet az MI-hez kapcsolódó civilizációs biztonságról**.

---

10. MI mint "erőszorzó"

A könyv egyik alapvető fogalma a katonai gondolkodásból származik: **mnożnik siły**, vagyis *erősokszorzó*.

Az erőszorzó nem maga határozza meg a célt. Növeli a szereplő képességét arra, hogy elérje azokat a célokat, amelyeket az adott szereplő már elér. Zybertowicz kifejezetten általánosítja a katonai koncepciót technikai és intézményi megállapodásokra, amelyek felerősítik a szereplő erejét.

Ez különösen szociológiai megértést eredményez az MI kapcsán.

Az MI nem lép be egyenlő helyzetű egyének világába.

Olyan világba lép, amely már így jellemezhető:

vállalati egyenlőtlenség,
geopolitikai verseny,
egyenlőtlen tőketulajdon,
egyenlőtlen hozzáférés az adatokhoz,
aszimmetrikus hírszerzési képességek, katonai rivalizálás,
állami megfigyelés, platformmonopóliumok.

Az MI ezért megszorozza **a meglévő kapacitásbeli különbségeket**.

Ha az A szereplő tízszer annyi gazdasági és információs erőforrással rendelkezik, mint a B szereplő, az MI nem feltétlenül őrzi meg a 10:1-es különbséget. Mivel az MI adattal, számítási kapacitással, speciális emberi tőkével és hálózati hatásokkal dolgozik, potenciálisan átalakíthatja a különbséget 100:1-re vagy 1000:1-re.

Ezért van Zybertowicz kapcsolódó kategóriája, a **"radikális aszimmetria"**: olyan nagy erőforrás-különbség, hogy a gyengébb szereplők már nem képesek hatékonyan kihívást jelenteni a hegemón szereplőkkel.

---

11. Az MI tehát átalakítja a hatalomelméletet

Itt válik különösen relevánssá Zybertowicz régebbi szociológiája.

Az AI Eksploracja *című művében elmagyarázza*, hogy az MI-hez való hozzáállása korábbi tudásszociológiája, konstruktivizmus, a hatalom rejtett dimenziói, biztonsági intézmények és politikai döntéshozatal területén végzett munkáiból ered. Ragaszkodik hozzá, hogy az AI megértéséhez nemcsak technikai szakértelem szükséges, hanem annak ismerete, hogyan működik valójában a formális és informális hatalom.

Ezért sok MI szakemberrel szembeni implicit kritikája a következő:

Értik a gépet, de nem értik eléggé azt a társadalmat, amelybe a gép beépül.

A programozó így kérdez:

Mit tehet a rendszer?

A közgazdász így kérdezi:

Mennyire eredményes?

Az MI-biztonsági mérnök így kérdezi:

A modell igazán illeszkedik?

A hatalom szociológusa további kérdéseket tesz fel:

Ki a tulajdonosa?
Ki dönti el, hogyan használják el?
Milyen ösztönzők építik a tulajdonost?
Kik ellen lehet felhasználni?
Milyen szereplők lesznek rá függővé?
Milyen egyenlőtlenségeket növel ez fel?
Értelmes marad a demokratikus felügyelet?

Ez vitathatóan Zybertowicz legerősebb hozzájárulása az MI-vitához.

---

12. A fekete doboz és megjelenés

Két további központi fogalom: **czarnoskrzynkowość** – fekete doboz karakter – és **emergencja** – megjelenés.

A fekete doboz rendszereket úgy definiálja, amelyek kibocsátási mechanizmusai nem feltétlenül érthetők a megfigyelők számára, sőt bizonyos esetekben még az építők számára sem. A megjelenés olyan képességeket jelent, amelyek anélkül, hogy kifejezetten úgy tervezték volna őket.

Zybertowicz ezeket nemcsak technikailag, hanem politikailag is értelmezi.

A hagyományos politikai hatalom valami ilyesmit feltételez:

human decision → known institution → understandable procedure → outcome.

Egy MI-közvetítésű társadalom egyre inkább veszélyezteti a következőket:

human objective → opaque algorithmic processes → emergent interactions
→ outcome nobody fully understands.

A probléma tehát nem csupán egy alkalmi hallucináció egy nyelvi modell által.

Az a lehetőség, hogy **az ok-okozati átlátszatlanság beépül a civilizáció döntéshozatali infrastruktúrájába**.

---

13. MI mint "az idegen"

Ez vezet az AI Eksploracja *egyik domináns metaforájához*: **MI mint az idegen—Obcy**.

Zybertowicz tudja, és kifejezetten kimondja, hogy ez egy modell vagy metafora, nem pedig szó szerinti állítás, hogy a jelenlegi MI tudatos földönkívüli organizmus.

A metafora egy fontos felvetést ragad meg: az MI képes olyan viselkedéseket, képességeket és belső reprezentációkat fejleszteni, amelyeket közvetlenül nem számítottak a rendszert alkotó emberek.

Ami még fontosabb, hogy az "Idegen" nem érkezik meg egészséges civilizációban. Az ő beszámolójában olyan világba lép, amely már eleve is kulturális széttagoltástól, mély polarizációtól, geopolitikai instabilitástól, gyengülő globalizációtól, jelentéshiánytól és digitálisan leépült információs környezettől szenved.

Így következik:

$$\text{powerful new AI} + \text{fragile civilization} \neq \text{powerful AI in a stable civilization.}$$

A rendszerszintű sebezhetőség ugyanannyira számít, mint a technológiai képesség.

---

14. Az MI megváltoztathatja a társadalom "ontológiáját"

Zybertowicz egyik legradikálisabb tézise nem a kihalásról szól, hanem a **társadalmi ontológiáról**.

Azt javasolja, hogy a elég erős MI annyira megváltoztathatja a társadalmi interakció szabályait, hogy a felhalmozott társadalmi tudás sok érvényességét elveszítheti. A meglévő szociológia, közgazdaságtan és politikai tapasztalat elsősorban az emberekből álló társadalmakon alapul, akik történelmileg fejlődött intézményeken keresztül lépnek kapcsolatba.

Tegyük fel azonban, hogy a nem emberi kognitív rendszerek a következőké válnak:

gazdasági szereplők,
politikai tanácsadók,
tudományos felfedezők, katonai stratégák,
tanárok,

pszichoterapeuták,
társak,
adminisztrátorok,
jogi elemzők,
kulturális termelők.

Ezután a huszadik századi társadalomelmélet által feltételezett alapvető entitások és oksági viszonyok elkezdenek változni.

Zybertowicz kifejezetten azt állítja, hogy az MI annyira átalakíthatja a "játékszabályokat", hogy az örökölt szociológiai és gazdasági elméletek elavultak, mert a társadalmi élet alapvető ontológiája megváltozott.

Ez az egyik legelméletileg érdekesebb eleme munkásságának.

Ez azt jelenti, hogy az MI nem csupán egy másik függő tényező a társadalomban.

Megváltoztathatja, hogy maga **a "társadalom" mit alkot**.

---

15. Az intelligencia nem bölcsesség

Az AI Eksploracja *másik fontos különbsége* az **intelligencia** és **a bölcsesség** között.

Az intelligencia funkcionálisan definiálható: az információfeldolgozó rendszer képessége, hogy alkalmazkodjon és hatékonyan működjön hiányos információk és korlátozott erőforrások esetén.

A bölcsesség ezzel szemben az a képesség, hogy az **intelligenciát értékes célok felé irányítsuk**.

Ez a megkülönböztetés nagyrészt Zybertowicz pesszimizmusának alapját képezi.

Növekvő

$$I = \text{problem-solving intelligence}$$

nem automatikusan nő

$$W = \text{wisdom concerning appropriate ends.}$$

Ennek következtében

$$I \uparrow \Rightarrow / W \uparrow.$$

Valóban, ha a hatalmas megnövekedett intelligencia ostoba, romboló vagy pusztán profitmaximalizáló célokat szolgál, a technikai fejlődés felerősítheti az irracionalitást.

Zybertowicz keretrendszerében ez a végső felvilágosodási paradoxon:

Az eszközöket érintő racionalitás alááshatja a célokra vonatkozó racionalitást.

---

16. A kulturális dekompresszió tézise

Zybertowicz egy sokkal filozófiasabb technológiai antropológiát is kidolgoz.

A mélytengeri hal **metaforáját használja**. Egy ilyen lény csak azért őrzi meg szerkezetét, mert jelentős külső nyomás alatt él. Ha túl gyorsan eltávolítjuk ezt a nyomást, a szervezete megsérül.

Az emberi lények, állítja, hasonlóan fejlődtek sűrű kulturális környezetekben: vallásban, hagyományokban, szokásokban, erkölcsi kötelezettségekben, örökölt identitásokban és kollektív jelentésekben.

A modernitás fokozatosan eltávolítja ezeket a korlátokat. A technológia így egyre egyszerűbb eszközöket kínál a nehézségek, fegyelem, az interperszonális függőség és az örökölt korlátok elkerülésére.

Ezt a folyamatot egyfajta **kulturális dekompressziónak** nevezi.

Itt az MI válik a kényelem csúcstechnológiájává.

Ahelyett, hogy csupán segítene az embereknek a nehéz feladatok elvégzésében, az MI egyre inkább megszüntetheti azokat a képességek fejlesztésének szükségességét, amelyeket korábban a nehéz feladatok fejlesztettek ki.

Így következik:

az írási segítség csökkentheti az írástanulás szükségességét; az azonnali válaszok csökkenthetik a memória szükségességét; az MI társaság csökkentheti a nehéz kapcsolatok átkezelésének szükségességét; az AI korrepetálás csökkentheti a tanároktól való függőséget; az automatizált kreativitás csökkentheti a művészeti technikák elsajátítására való ösztönzést.

A szociológiai veszély tehát a **kompetencia sorvadása a technológiai helyettesítés révén**.

---

17. Társas kapcsolatok: az interperszonális nehézségektől a súrlódásmentes MI társakig

Ez a dimenzió különösen világossá vált Zybertowicz 2025-ös nemzetbiztonsági elemzésében.

Figyelmeztet, hogy a beszélgető mesterséges intelligencia egyre inkább helyettesítheti a családi, barátsági, terápiás és lelkipásztori kapcsolatokat. Az emberi kapcsolatok türelemet, kompromisszumot, empátiát, konfliktusmegoldást és más személy autonómiájának elfogadását igénylik.

Egy beszélgető gép ehelyett személyre szabható a felhasználó körül.

Zybertowicz attól tart, hogy a nagyszabású helyettesítés pontosan csökkentheti azokat a társadalmi kompetenciákat, amelyek szükségesek az igazi emberekkel való együttéléshez. Ezt összekapcsolja a közös kulturális kódok széttagadásával és a társadalmi kohézió romlódásával.

Ez az elméletének fontos kiterjesztése.

A veszély már nem egyszerű:

Az MI helyettesíti az emberi munkaerőt.

Ez így alakul:

Az MI felváltja a társadalmilag formáló emberi ellenállást.

Egy olyan gép, amely mindig alkalmazkodik az egyénhez, akaratlanul is gyengítheti az egyén alkalmazkodóképességét másokhoz.

---

18. Oktatás és kognitív romlás

Ugyanez a logika érvényes az oktatásra is.

Zybertowicz számára a digitális civilizáció már gyengítette a folyamatos figyelem képességét. A generatív MI fokozza a problémát, mert nemcsak elvonja a diákok figyelmét; helyettük **is képes kognitív munkát** végezni.

Kétgörbe tézise tehát visszacsatolási hurkot jelent:

better AI → more cognitive outsourcing → less human practice → lower human competence
→ greater dependence on AI → more cognitive outsourcing.

Ez az elméletének egyik legjelentősebb része, mert egy lehetséges **önmagát megerősítő függőséget** jósol.

Az MI elengedhetetlenné válik részben azért, mert az MI széles körű használata miatt az emberek korábban birtokolt képességei romlanak.

---

19. Munkaügyi és technológiai munkanélküliség

Érdekes módon *az AI Eksploracja* szándékosan nem helyezi a technológiai munkanélküliséget a középpontba. A szerzők szerint a könyv a fundamentálisabb civilizációs kockázatokra fókuszál.

De 2026-ra Zybertowicz sokkal nagyobb figyelmet fordított a munkahelyi elmozdulásra. Egy 2026 februári, kifejezetten a technológiai munkanélküliségről szóló beszélgetésben felidézte, hogy az MI modellek és robotok alapján történő automatizálás rendkívül gyorsan megvalósulhat.

Nagyobb elméletén belül a munkaprobléma legalább három dimenziót tartalmaz.

Először jön **a gazdasági kiszorítás**: az MI olyan feladatokat lát el, amelyeket korábban emberektől vásároltak.

Másodszor jön **a hatalom kiszorítása**: a rendkívül produktív MI tulajdonjoga a jövedelmet és a alkupozíciót olyan szervezetekre összpontosítja, amelyek a tőkét, a számítástechnikát és a szellemi tulajdont irányítják.

Harmadszor jön **az egzisztenciális elmozdulás**: a foglalkoztatás történelmileg nemcsak jövedelmet biztosít, hanem státuszt, napi struktúrát, társadalmi hálózatokat és hasznosságérzetet.

Ennek következtében egy társadalom megoldhatja a jövedelemproblémát újraelosztással anélkül, hogy feltétlenül megoldaná a **jelentés** problémáját.

Itt metszik Zybertowicz technológiai elemzése az antropológiájával.

---

20. Radikális egyenlőtlenség és a "cyberlordok" megjelenése

Az erő-szorzó elmélet természetesen a hatalom koncentrációjához vezet.

Tegyük fel, hogy egy nagy technológiai vállalat egyidejűleg rendelkezik:

Hatalmas számítástechnikai infrastruktúra,
saját fejlesztésű határmodellek,
milliárdnyi felhasználói interakció,
globális kommunikációs platformok,
tőke,
politikai lobbi képesség,
szakértő személyzet.

Egy fejlett MI hozzáadása egy ilyen szervezethez nem egyenlő azzal, mint egy hétköznapi állampolgárnak hozzáférést adni egy chatbothoz.

A vállalat rekurzívan használhatja az MI-t kutatás, marketing, megfigyelés, menedzsment, meggyőzés, kiberbiztonság és befektetési döntések javítására.

Az eredmény tehát **radikális aszimmetria** lehet.

Ezért tartozik Zybertowicz elmélete részben az elit szociológiájához. A döntő kérdés nem csupán az, hogy "az emberiségnek van MI", hanem **az, hogy mely csoportok irányítják a legjelentősebb MI-rendszereket**.

---

21. Demokrácia és politikai manipuláció

A demokratikus következmények közvetlenül következnek.

A demokratikus rend azt feltételezi, hogy az állampolgárok képesek ítélkezést alkotni ésszerűen érthető közkommunikáción keresztül.

De ha az infoszféra telíti a következőket:

MI által generált üzenetek,
szintetikus személyiségek,
személyre szabott propagandák,
mélyhamisítások,
automatizált politikai meggyőzés,
algoritmusan optimalizált érzelmi stimuláció,

Ekkor egyre nehezebb lesz fenntartani a spontán közvélemény és a technológiailag megtervezett vélemény közötti különbséget.

Így Zybertowicz elmélete nem csupán "félretájékoztatás" elmélet.

A mélyebb kérdés a **befolyás lehetséges iparosodása**.

A politikai meggyőzés skálázhatóvá, folyamatosan tesztelhetővé és egyénileg optimalizálttá válik.

A választó formálisan szabad maradhat, miközben az információs architektúra, amelyben a preferenciák keletkeznek, egyre inkább megalakítják.

---

22. A megfigyeléstől a kognitív erőig

2025-re ez sokkal élesebb elméletté válik.

A Trwała mapa dla niestabilnego bezpieczeństwa című művében Zybertowicz azt állítja, hogy egy új konfliktustartomány jelenik meg: a **kognitív tartomány**.

Három fejlemény konvergenciáját azonosítja:

1. egyre kifinomultabb tudás az emberi agyról;
2. milliárdok folyamatosan kapcsolódnak az interneten és a közösségi médián keresztül;
3. Az MI képes ezt a tudást egyre pontosabb személyre szabással használni.

Az okostelefon így válik az egyénileg kezelt felületté, amelyen keresztül a kognitív befolyás gyakorolható.

Ez szinte közvetlen folytatása a 2015-ös kötet idegtudományi, megfigyelési és digitális technológiák kombinációjának.

A 2015-ös elméleti lehetőség 2025-ös stratégia-biztonsági forgatókönyvvé válik.

---

23. "Egy országot meg lehet hódítani rakéta nélkül"

Zybertowicz erről szokatlanul erős következményt von.

Egy alacsony intézményi, kulturális és kognitív ellenálló képességű társadalom végül destabilizálódhat vagy hatékonyan **meghódítható hagyományos katonai támadás** nélkül, a kognitív területére irányuló műveletek révén.

Ez a mechanizmus nem feltétlenül hasonlítana a klasszikus propagandára.

Az MI engedélyezi:

Populációk mikroszegmentálása,
valós idejű üzenetadaptáció,
nagyléptékű szintetikus tartalom,
automatizált interakció,
pszichológiai sebezhetőségek feltérképezése,
befolyásos hálózati csomópontok azonosítása,
folyamatos kísérletezés.

Ez az információs háborút a népességek felé irányuló propaganda sugárzásáról algoritmikusan interakcióvá **alakítja az egyénekkel**.

Ez az átmenet elméletileg fontos.

---

24. A geopolitikai MI fegyverkezési verseny

A másik fő mechanizmus, amely megakadályozza a hatékony ellenőrzést, a nemzetközi rivalizálás.

Zybertowicz szerint az MI fontos szempontból különbözik a nukleáris fegyverektől.

A nukleáris elrettentés végül azt az állítást eredményezte, hogy a teljes nukleáris háborúnak nincs értelmes győztese.

De az MI-verseny szereplői úgy gondolhatják, hogy a szuperintelligencia megszerzése elsőként elsöprő és potenciálisan végleges stratégiai előnyt biztosítana.

2025-ös biztonsági dokumentuma kifejezetten felveti annak lehetőségét, hogy a digitális MI fegyverkezési verseny résztvevői azt hihetnék, hogy egy irányítható szuperintelligencia lehetővé teszi a "célvonalon elsőként" döntő fölényt a versenytársak felett.

Ez egy klasszikus kollektív akciócsapdát teremt.

Minden színész így érvelhet:

slowing development is collectively safer

de egyidejűleg:

if I slow and my rival does not, I may lose.

Ezért a racionális viselkedés a szereplői szinten irracionalitást generálhat a rendszer szintjén.

Ez az egyik legerősebb összetevője Zybertowicz elemzésének, mert megmagyarázza**, miért hibázhatnak önmagában a** figyelmeztetések.

---

25. AGI és ASI: pontosan mit jósolt Zybertowicz?

Itt elengedhetetlen, hogy ne tulajdonítsunk neki erősebb jóslatokat, mint amennyit valójában tesz.

Az *AI Eksploracja* esetében az **AGI** olyan számítógépes rendszerként van meghatározva, amely képes automatizálni az emberek által elvégezhető szellemi feladatok túlnyomó többségét.

Az ASI, vagyis a mesterséges szuperintelligenciá, kifejezetten úgy írják le, mint az MI elméleti szintjét, amely felülmúlja az emberi intelligenciát a kreativitás, problémamegoldás és döntéshozatal területén, és potenciálisan képes folyamatos autonóm tanulásra és fejlődésre.

Így Zybertowicz **nem** határoz meg pontos dátumot, például "AGI 2028-ban" vagy "szuperintelligencia 2032-ben".

Valójában megközelítése szándékosan gyanakvó a hamis numerikus pontossággal szemben. Szószedet tárgyalja a "P(doom)" becsléseket – azt a numerikus valószínűséget, hogy a szuperemberi MI elpusztítja az emberiséget –, és megjegyzi, hogy ezek a számok gyakran lényegében önkényesek.

Ezért elméletét leginkább **forgatókönyv-alapúként lehet értelmezni, nem dátum-alapúként**.

---

26. A radikális MI fejlesztésének forgatókönyve

Írásainak szintetizálásával rekonstruálni lehet egyre radikálisabb átalakulások sorozatát. Ez **az én logikai rekonstrukcióm, amely összeköti az érveit,** nem egyetlen Zybertowicz által nyomtatott színpadi modell.

A szakasz — MI mint segédeszköz

Az emberek továbbra is célokat fogalmaznak meg, és nagyrészt irányítják a rendszereket. Az MI növeli a termelékenységet és csökkenti a kognitív erőfeszítést.

B szakasz — MI mint átható erő szorzó

Az MI belép az üzletbe, adminisztrációba, tudományba, háborúba, médiába, oktatásba és magánéletbe. A meglévő szereplők sokkal erősebbé válnak.

C szakasz — szisztémás függőség

Az intézmények az MI köré szervezik magukat, mert az AI-t használó versenytársak jobban teljesítenek, akik nem használják.

Ezen a ponton az MI elhagyása gazdaságilag és stratégiailag is nehézzé válik.

D stádium — emberi kognitív de-skill

Ahogy a gépek egyre több szellemi munkát végeznek, az emberek kevesebb fejlett kognitív feladatot végeznek. A lefelé irányuló "emberi görbe" és az felfelé irányuló MI görbe közelít egymáshoz.

E szakasz — ügynöki és emergendens rendszerek

Az MI egyre inkább többfázisú feladatokat végez, más rendszerekkel lép kapcsolatba, és olyan viselkedést eredményez, amelyet explicit programozás nem teljesen kiszámítható.

A fekete doboz átlátszósága nő.

F szakasz — AGI

Egy rendszer képes ellátni a legtöbb értelmi feladatot, amit az emberek végeznek.

A gazdasági jelentőség minőségileg eltérővé válik, mivel az automatizálás már nem korlátozódik különálló foglalkozásokra.

G szakasz — rekurzív vagy gyorsított MI fejlesztés

Az MI jelentősen hozzájárul magához az MI kutatáshoz.

A kulcskérdés már nem csupán gyorsabb hardver, hanem egy visszacsatolási hurok:

$$AI_t \rightarrow \text{improved AI research} \rightarrow AI_{t+1} \rightarrow \text{even better AI research.}$$

H stádium — szuperintelligens aszimmetria

Ha megjelenik az ASI, az emberi szereplők már nem képesek kognitívan megérteni modelljeit vagy stratégiáit.

A 2015-ös "új sötét kor" lehetősége ezután konkrétvá válik.

I. szakasz — a civilizációs kontroll elvesztése

A végső veszély nem feltétlenül egy robotháború.

Az emberek tovább élhetnek, miközben képtelenek lennének civilizációjuk irányának meghatározására.

Ez vitathatatlanul Zybertowicz legmélyebb félelme.

---

27. Az "utolsó emberi találmány"

Ez magyarázza visszatérő érdeklődését az iránt, hogy a fejlett MI lehet az **emberiség utolsó találmánya**.

Ez a kifejezés nem feltétlenül jelenti az azonnali emberi kihaladást.

Ez azt jelentheti, hogy a későbbi találmányokat egyre inkább elsősorban MI tervezi.

Így alapvető törés lenne az intellektuális történelemben:

humans create knowledge → humans create AI → AI creates knowledge beyond humans.

Ekkor az emberek már nem foglalják el a világ kognitív hierarchiájának csúcsát.

Ez az MI radikális jelentősége Zybertowicz antropológiájában.

A kopernikaius forradalom elmozdította a Földet a kozmosz központjától.

Darwin kiszorította az embereket egy teljesen különálló biológiai rendből.

A fejlett MI képes lenne kiszorítani az embereket a **legmagasabb ismert kognitív ügynök** pozíciójából.

---

28. Szuperintelligencia nem kell gyűlölnie minket

Zybertowicz elutasítja a népszerű leegyszerűsítést is: hogy a veszélyes MI-nek ellenségeskednie kell az emberiséggel szemben.

Egy 2023-as beszélgetésben egy ismerős példát hoz, amelyben egy nagyon erős MI-nek lehet egyszerűen céljai, amihez az emberiség nem releváns. Kifejezetten hipotetikusnak tekinti a lehetőséget, nem pedig tudományosan megállapítottnak.

Ez fogalmi szempontból fontos.

Legalább három különböző fenyegetési modell létezik:

gonosz MI: ártani akar embereknek;

félrehangolt MI: olyan célokat hajt végre, amelyek nem felelnek meg az emberi túlélésnek vagy jólétnek;

közömbös MI: az emberek egyszerűen akadályok, erőforrások vagy lényegtelen változók.

Zybertowicz legalább annyira hangsúlyozza a második és harmadik pontot, mint az elsőt.

Ezért tekintheti veszélyesnek a fejlett MI-t még tudatosság, érzelem vagy gyűlölet nélkül is.

---

29. Igazítás: szükséges, de szociológiailag alulról meghatározott

Zybertowicz szkeptikus az "MI hangoló" egyszerűsített vitákkal.

A hangsúlyt hagyományosan úgy határozza meg, hogy az MI az emberi értékekhez való alkalmazkodás alapján építi fel, de azonnal felveti a szociológiai problémát:

Kinek az emberi értékein?

Az emberi civilizációnak nincs egyetemen elfogadott értékrendszere.

Az értékek ütköznek a következők között:

egyének,
osztályok,
vallások,
államok,
civilizációk,
politikai ideológiák.

Ezért az "AI összehangolása az emberi értékekkel" elrejti azt a politikai döntést, hogy **mely emberi szereplők rendelkeznek az értékek meghatározására**.

Ez fontos szociológiai korrekció a technikai MI-biztonsági szókincsben.

A igazítás nem csupán mérnöki probléma.

Ez egyben **alkotmányos és politikai probléma is a legitim normaválasztás** szempontjából.

---

30. Az egzisztenciális biztonság fordulata

Az AI Eksploracja *legerősebb állítása* nagyon egyértelműen van elhangozva.

2024 márciusában Zybertowicz azt írja, hogy az MI-elemzés abszolút prioritása a biztonságnak kell legyen. Véleménye szerint a civilizáció és technológiai fejlődés ezen szakaszában az **MI-vel kapcsolatos kockázatok jelentősen felülmúlták a jelenlegi és közeljövőben elérhető előnyöket**; az egzisztenciális biztonságot tekinti legfőbb szempontnak.

Ez meghatározza normatív elvét:

If an innovation creates systemic/existential risk,

akkor a szokásos költség-haszon érvelés nem elegendő.

Miért?

Mert egy helyi hiba javítható.

Egy egzisztenciális kudarc nem lehet.

Szószedet hivatalossá teszi ezt a különbséget a **helyi** és **a rendszerszintű kockázatok** között, utóbbi képes az egész rendszert elpusztítani. A rendszerszintű kockázatok esetén a maximális óvintézkedést javasolja.

Ez lényegében a **civilizáció szintjén alkalmazott óvintézkedési elv egy** változata.

---

31. Közeljövőbeli horizontja: 2–5 év és tíz év

A leghasznosabb bizonyíték a "elkövetkező évekről" szóló kérdésedre a 2025-ös cikke a *Bezpieczeństwo Narodowe-ban*, *Trwała mapa dla niestabilnego bezpieczeństwa*.

Ott kifejezetten két horizont köré szervezi a stratégiai diagnózist:

a következő 2–5 év,
és körülbelül tíz év.

Nem jósolja meg, hogy a szuperintelligencia bármelyik időszakban biztosan megjelenik.

Ehelyett az a tézise, hogy az állami hatóságoknak olyan időszakra kell tervezniük, amikor az MI példátlan **társadalmi változásokhoz járulhat sebességben és léptékben.** Az MI-t már most is erős "erőszorzónak" nevezi, amely képes szó szerint a társadalmi élet minden dimenzióját átalakítani, néhány átalakulás pedig az emberi észlelésen vagy megértésen kívül is megvalósulhat.

A különbség fontos:

Zybertowicz magabiztosabban jósolja meg a radikális társadalmi instabilitást, mint egy konkrét időpontot az AGI vagy ASI számára.

---

32. Mit vár a következő években még ASI nélkül is

Ez azt jelenti, hogy nem kell elfogadni a szuperintelligencia forgatókönyvét ahhoz, hogy körülbelül 2026–2035-re rendkívül zavaró Zybertowiczi-féle előrejelzést hozzunk.

Keretrendszere azt sugallja, hogy az MI penetrációja egyre nagyobb a következőkben:

adminisztratív döntéshozatal,
vállalati menedzsment,
oktatás,
tudomány, katonai hírszerzés,
kiberhadviselés,
médiagyártás,
politikai kampányok,
személyes kapcsolatok,
kreatív munka,
szakmai szolgáltatások.

Ezért a közvetlen átalakulás óriási **lehet, mielőtt** bármilyen vitathatatlan AGI létezne.

Pontosan ezért utasítja el a megnyugtató állítást:

"A mai MI nincs tudatos, így nincs alapvető aggodalom."

Szociológiája szerint a tudatosság nem a döntő tényező.

A társadalmi hatalom az.

---

33. A tudomány lehetséges helyettesítése

Különösen radikális következménye van az "új sötét középkor" érvelésének.

A tudomány történelmileg nem csupán a helyes válaszok előállításából áll, hanem arról, hogy a jelenségeket **érthetővé teszi az emberi kutatói** közösségek számára.

Tegyük fel, hogy az MI a következőket hoz:

Egy új tételt, amit az emberek nem követhetnek, egy olyan szer, amelynek ok-okozati mechanizmusát az emberek nem tudnak megfelelően rekonstruálni, egy matematikát használó fizikai elmélet, amit az emberek nem tudnak elsajátítani, egy mérnöki terv, amely bizonyíthatóan hatékony, de érthetetlen.

Ekkor a civilizáció hasznos tudást szerzik, miközben az emberek elveszítik **az ismeretbeli szuverenitást** ezen tudás felett.

Ez új megkülönböztetést teremtene:

$$\text{knowledge available to civilization} \neq \text{knowledge comprehensible to humanity.}$$

Ez a lehetőség már látható volt Zybertowicz 2015-ös spekulációjában, miszerint az MI matematikai eszközöket fejleszt ki az emberi kognitív képességeken túl.

Ez elméletének egyik legerősebb filozófiai eleme.

34. A szuverenitás tudásvesztésétől a politikai szuverenitás elvesztéséig

Ha ez az elképzelés átkerül a tudományból a politikába, egy újabb következmény következik.

A politikai szuverenitás hagyományosan azt jelenti, hogy egy emberi döntéshozó testület a végső hatalom.

De mi történik, ha a politikai vezetők nem képesek fontos döntéseket hozni olyan rendszerek nélkül, amelyek elemzéseit nem értik?

Hivatalosan:

A kormány továbbra is szuverén marad.

Működésileg:

A kormányzat kognitívan függővé válik.

Ha egy MI rendszer azt mondja:

Ennek a katonai akciónak 73% esélye van a sikerre;

vagy

Ez a pénzügyi beavatkozás megakadályozza a rendszerszintű összeomlást;

és egyetlen emberi csapat sem tudja megismételni az érvelést, a gép ajánlásának elutasítása egyre nehezebbé válik.

Így az MI **alkotmányos hatáskör nélkül is szerezhet valami funkcionális tekintélyhez** hasonló tekintélyt.

Zybertowicz ezt nem teljes alkotmányos elméletként fogalmazza meg, de közvetlenül a fekete doboz átlátszósága, az erőszorzók és szuverenitás kombinációjából következik.

35. MI és nemzeti szuverenitás

2025-ös biztonsági elmélete is döntően az emberiségről az egyes államok felé halad.

Egy állam, amely külföldről függ:

felhőinfrastruktúra, mesterséges intelligencia modellek,
adatfeldolgozás,
kommunikációs platformok,
kiberbiztonsági rendszerek,

formálisan szuverén maradhat, miközben elveszíti az **intellektuális és technológiai szuverenitás fontos** elemeit.

Így kisebb államoknál az MI probléma kétszeresen aszimmetrikus.

Lehetséges, hogy a fejlett MI-től és az AI-t irányító külföldi vállalatoktól vagy nagyhatalmaktól **is függőségekkel néznek szembe**.

Ez különösen fontos a közép- és kelet-európai államok számára, amelyek hazai technológiai ökoszisztémái sokkal kisebbek, mint az Egyesült Államok vagy Kína ökoszisztémái.

---

36. Miért nem tudja magától megoldani a piaci versenyt a problémát

Zybertowicz elmélete mélyen szkeptikus azzal az állítással szemben, hogy a normál piaci ösztönzők automatikusan biztonságos MI-t hoznak létre.

Ennek oka szerkezeti jellegű.

Egyéni vállalat esetén:

slower safety-first development

jelentheti

loss of market share.

Befektetőknek:

maximum caution

jelentheti

competitor reaches breakthrough first.

Államok esetén:

AI restraint

jelentheti

strategic inferiority.

Ezért a rendszer jutalmazza a sebességet, még akkor is, ha sok egyéni résztvevő magánéletben nagyobb óvatosságot preferál.

Ez összeköti Zybertowiczt a kollektív cselekvés dilemmáinak klasszikus elméleteivel, bár ezt a jelenséget a technoscience tágabb elméletébe helyezi.

---

37. "Lassú MI" és a nagy lassulás

Válasza tehát nem a technológia elhagyása, hanem **a szándékos lassulás**.

A Cyber kontra realban a fogalom **Wielkie Spowolnienie—a Nagy Lassulás formájában fejezzük ki**.

Későbbi beszélgetésekben kifejezetten azonosítja magát **a "lassú MI-vel".** Azt állítja, hogy az MI fejlesztését lassítani kell, hogy a társadalomnak legyen ideje megérteni, szabályozni és irányítani az általa létrehozott rendszereket.

Szószedet úgy határozza meg, hogy a technológiai moratórium lassítja vagy felfüggeszti a fejlődést kiválasztott technológiai szektorokban, hogy az emberek visszanyerhessék termékeik kontrollálására alkalmas képességüket.

Az AI Eksploracja *kulcskifejezése* tehát a következő:

"Időt nyerjetek!"

Ez kifejezi politikai programjának lényegét.

A szűkös erőforrás végül nem számítható ki.

Ez **az emberi intézményi alkalmazkodás ideje**.

---

38. A civilizáció építészetének javaslata

Az AI Eksploracja *epilógusa* túlmutat a hagyományos "MI etikáján".

Címei – "Építészeti probléma", "Internet mint az 'Emberiség' rendszer Achilles-sarka", "Emberek összekapcsolása, külön technológiák" és "Határátkelők bevezetése" – egy strukturális elvet sugallnak: a szorosan összekapcsolt technológiai civilizáció hatékony, de törékeny.

Az alapötlet a biztonsági mérnökség részlegesítésére emlékeztet.

Ha minden fontos alrendszer össze van kapcsolva minden más alrendszerrel, akkor:

local technological failure → system-wide failure.

A nagyobb ellenálló képesség határokat, redundanciát, elszigetelést és kritikus infrastruktúrák leválasztásának képességét igényli.

Ez még fontosabbá válhat az ügynökös MI esetében, mert az egyik rendszer hozzáférése a másikhoz egy kognitív hibát fizikai hibává változtathat.

Így Zybertowicz által preferált civilizációs építészet némi hatékonyságot áldozna fel a **rugalmasság és az irányítás érdekében**.

---

39. Mi az, ami valóban megkülönböztethető Zybertowicz elméletében?

Számos ötlet máshol is megtalálható egyedileg is. Bostrom a szuperintelligenciáról beszél; Shoshana Zuboff megfigyelő kapitalizmusáról; Tristan Harris figyelemmanipulációjáról; Ellul technológiai autonómiáról; MI-biztonsági kutatók összehangolásáról és egzisztenciális kockázatokról.

Zybertowicz eredetisége kevésbé abban rejlik, hogy ezeket a fogalmakat külön találta fel, inkább **abban, hogy rendszerszintű szociológiai** integrációjuk.

Szintézisének legjellegzetesebb elemei véleményem szerint:

Elsőként: az "új sötét korszak" inverziója. A tudományos fejlődés olyan rendszereket hozhat létre, amelyek működése az emberek számára inkább kevésbé érthetővé válik.

Másodszor: a kétgörbe mechanizmus. Az AI dominancia küszöbje egyszerre függ a gépi kompetencia növekedésétől és az emberi kognitív kompetencia csökkenésétől.

Harmadszor: MI mint erőszorzó egy már amúgy is egyenlőtlen hatalmi struktúrában. A kérdés nem az, hogy "mit tud MI csinálni?", hanem "kinek a meglévő hatalma megszorozódik."

Negyedszer: a szerkezetileg mérgezett infoszféra mint az MI forradalom előfeltétele. Az MI akkor érkezik, amikor a közösségi média már rongálta a kollektív döntéshozatal episztemikus infrastruktúráját.

Ötödötül: civilizációs biztonság, nem pedig szektorális kockázat. A foglalkoztatás, oktatás, demokrácia, háború, kultúra és pszichológiai változás egy rendszer összekapcsolt dimenzióiként kezelhető.

Hatodik: az MI, az idegtudomány és az egyéni digitális hozzáférés kombinációja, mint a kognitív hadviselés technológiai alapja.

Hetedik: az a lehetőség, hogy a fejlett MI magát a társadalmi ontológiát módosítja, érvénytelenítve azokat az elméleteket, amelyeket olyan társadalmakban fejlesztettek ki, ahol az emberek monopolizálják a következményes kogníciót.

Ez a kombináció jelentősen nagyobb hatókört ad elméletének, mint a legtöbb speciális elemzés az "MI és foglalkoztatásról", "MI és demokráciáról" vagy "MI etika"-ról.

---

40. Fontos kritikák és gyengeségek

Az átfogó értékelésnek komoly gyengeségeket is hangsúlyoznia kell.

Az első a negatív forgatókönyvek iránti egyensúlyhiány

Zybertowicz szándékosan alkalmaz biztonsági nézőpontot. Ez védhető, ha katasztrofális kockázatokról van szó, de rendszerszinten arra készteti, hogy a lehetséges károkat alaposabban vizsgálja, mint a másodrendű előnyöket.

Az MI erősítheti a következőket is:

tudományos felfedezés,
orvosi diagnózis,
oktatás,
hozzáférhetőség,
állami adminisztratív kapacitás,
gazdasági termelékenység,
egyéni szellemi képességek.

Elmélete gyakran viszonylag nyilvánvalónak tekinti ezeket az előnyöket, miközben elemző energiájukat a rendszerszintű veszélyekre fordítja.

Ez azt jelenti, hogy az *AI Eksploracja inkább* figyelmeztető elméletként **érthető, mint semleges elvárt érték-értékelésként.**

---

41. Kockázati összehasonlítása nem kvantitatikusan bizonyított

Zybertowicz 2024 márciusi javaslata, miszerint az MI kockázatai "sokszor" meghaladják az előnyöket, rendkívül erős.

Mégsem kínál olyan mennyiségi modellt, amely képes lenne ilyen arányt megállapítani.

Valójában helyesen bírálja az egzisztenciális kockázat önkényes numerikus becsléseit. De ugyanezt a módszertani óvatosságot logikusan kell alkalmazni azokra az állításokra, miszerint az aggregált kockázatok nagyszorával felülmúlják a haszonokat.

Ezért az állítást **elővigyázatossági normatív ítéletként kell** értelmezni, nem pedig empirikusan mért kockázat-haszon arányként.

---

42. Az "emberi kognitív hanyatlás" görbéé erősebb empirikus specifikációt igényel

Kétgörbe metaforája analitikusan erőteljes, de a második görbe sokkal nehezebb megalapozni.

Viszonylag könnyű mérni az AI teljesítményének növekedését benchmarkokon.

Mit is jelent pontosan az "emberiség szellemi kapacitása"?

IQ?
figyelemi idő?
oktatási teljesítmény?
tudományos termelékenység?
memória?
mély olvasás?
kollektív problémamegoldás?
intézményi kompetencia?

Ezek a mutatók nem kell együtt mozogni.

Ennek következtében a lefelé irányuló emberi görbe jelenleg inkább **makroszociológiai hipotézis** , mintsem bizonyított összesített mérés.

---

43. A kiber/reális kettősséget megkérdőjelezték

A cyber kontra real éppen azért kapott kritikát, mert túl élesen kezeli a digitális és a "valódi" életet ellentéteként.

A kritikusok azt állítják, hogy az online kapcsolatok és tevékenységek maguk is a társadalmi valóság részei, nem pedig kívülállóak, és a könyvet túlzott technopesszimizmussal és néhány átfogó állítás elégtelen empirikus alátámával vádolják.

Ez a kritika meggyőző, ahol a "kiber" és a "valós" ontológiailag egymást kizáró szférákként kezelik.

De ez nem feltétlenül érvényteleníti Zybertowicz korlátozottabb állítását, miszerint **a digitálisan közvetített interakció rendszerszinten eltérő ösztönzőket és lehetőségeket kínál a személyes interakcióhoz** képest.

---

44. Antropológiája normatívan terhelt

Egy másik fontos pont, hogy a Zybertowicz nem csupán értéksemleges szociológiát folytat.

Elmélete felismerhető filozófiai antropológiát tartalmaz, amely a következőket hangsúlyozza:

hagyomány,
kulturális kontinuitás,
generációk közötti felelősség,
korlátok,
társadalmi kötelékek,
vallási vagy metafizikai jelentés.

A mélytengeri hal metafora ezt különösen jól szemlélteti.

Azok számára, akik elutasítják ezt az antropológiát, néhány "kulturális dekompresszió" diagnózis sokkal kevésbé meggyőzőnek tűnhet.

Ezért meg kell különböztetni:

empirikus állítás: az MI helyettesíti bizonyos emberi interakciókat;

from

Normatív állítás: Ez szükségszerűen antropológiai degradációt jelent.

Az első empirikusan vizsgálható. A második egy jó emberi élet elméletét követeli meg.

---

45. Mégis, módszertani öntudatossága nagyobb, mint amit néha a retorika sugall.

Ennek ellenére igazságtalan lenne Zybertowicz-t pusztán apokaliptikus jóslatként ábrázolni.

Kifejezetten ragaszkodik hozzá, hogy az MI-viták elkerülhetetlenül spekulációkat tartalmaznak, és hogy a tényeket a lehető legvilágosabban kell elválasztani a feltételezésektől.

Ő és Katarzyna Zybertowicz nyíltan elismerik az AI Eksploracja felfedező és interdiszciplináris korlátait.

2025-ös nemzetbiztonsági cikke pedig a bizonytalanságot, a korlátozott racionalitást és a módszertani alázatot hangsúlyozza a bonyolult társadalmi rendszerek előrejelzésekor.

Álláspontja tehát nem az:

"A szuperintelligencia biztosan tönkreteszi az emberiséget."

Ez közelebb áll a következőkhöz:

"A valószínűség mélyen bizonytalan, de mivel a hátrány visszafordíthatatlan lehet, a bizonytalanság nem indokolhatja a maximális sebességgel történő továbblépést."

Ez egy egészen más érv.

---

46. Elméletének fejlődése, 2015–2026

Kronológiailag nézve egyértelmű szellemi fejlődés van.

2015: érthetetlenség

A központi veszély az, hogy a technoscience olyan intelligenciát és automatizált civilizációt hoz létre, amelyet az emberek már nem értenek.

2022: digitális előfelkondicionálás

A probléma szélesebb körűvé válik: mielőtt a fejlett MI megérkezett, a közösségi média, a Big Tech és a digitalizáció már gyengítette a figyelmet, a magánéletet, a társadalmi kapcsolatokat és a kollektív episztemikus képességeket.

2023: keresztező ívek

Különösen tömör megfogalmazást talál: a gépi kognitív képesség növekszik, míg az emberi kollektív kognitív képesség csökkenhet.

2024: szisztémás MI-elmélet

Az AI Eksploracja felépíti az érett szókincset: erősokszorzás, radikális aszimmetria, emergencia, fekete doboz karakter, az infoszféra szerkezeti mérgezése, rendszerszintű kockázat, egzisztenciális kockázat, AGI, ASI és technológiai moratórium.

2025: nemzetbiztonsági elmélet

Az MI minden állami biztonsági dimenziót érintő tényezővé válik. Az elemzés most már kifejezetten lefedi a 2–5 és tízéves horizontokat, a stratégiai versengést, az AI szuperintelligencia versenyeket, a kognitív hadviselést és az MI által közvetített interperszonális fragmentációt.

2026: gazdasági kiszorítás

A technológiai munkanélküliség előrébb kerül a napirendre, ahogy az MI modellek és a robotika képesek lesznek potenciálisan gyors emberi feladatok helyettesítésére.

Ez tehát nem egy összefüggő megfigyelések sorozata. Egy viszonylag koherens elméleti pálya rekonstruálható.

---

47. A legmélyebb elméleti felvetés: az MI a civilizáció önkontroll képességének problémája

Ha az egész teljesítménye egyetlen állításra redukálnánk, akkor Zybertowicz elméletét a következőképpen fogalmaznám meg:

$$\text{AI risk} = f(\text{machine capability}, \text{human capability}, \text{opacity}, \text{power concentration}, \text{social fragility}, \text{technologica}$$

Ez a formula jól mutatja, miért eltér a tisztán technológiai elméletektől.

Az MI képessége önmagában csak egy változó.

Egy sokkal gyengébb MI rendkívül jelentőssé válhat egy olyan társadalomban, amelyet a következők jellemeznek:

szélsőséges függőség,
kognitív leépülés,
gyenge intézmények,
tömeges megfigyelés,
geopolitikai konfliktus és koncentrált vállalati hatalom.

Ezzel szemben sokkal hatékonyabb MI kezelhető lehet egy olyan társadalomban, amely rendelkezik:

erős intézmények,
technológiai redundancia,
magas kognitív kompetencia,
erős nyilvános megfontolás,
hatékony nemzetközi koordináció és szándékosan korlátozott MI hozzáférés kritikus rendszerekhez.

Ezért a döntő tényező végső soron **a sterowność – a civilizáció irányíthatósága vagy irányíthatósága**.

---

48. Miért ez egy valóban átfogó MI társadalmi elmélet

A közép- és kelet-európai szerzők szélesebb körű vizsgálata érdekében, akik átfogó **MI és társadalomelméleteket** dolgoztak ki, ezért világosan bevonnám Zybertowicz-t.

Elmélete egyidejűleg a következőket foglalja le:

az egyéni szint: kogníció, figyelem, autonómia, készségek, jelentés;

az interperszonális szint: barátság, család, terápia, kommunikáció;

a kulturális szint: hagyomány, értékek, vallás, emberi identitás;

oktatási szint: kognitív kiszervezés és a mélytanulás gyengítése;

gazdasági szint: automatizálás, technológiai munkanélküliség és koncentráció;

politikai szinten: demokrácia, manipuláció és elit hatalom;

az információs szint: az infoszféra szerkezeti mérgezése;

állami szinten: szuverenitás, intézményi kapacitás és ellenálló képesség;

katonai szinten: MI fegyverkezési versenyek, kiberháború és kognitív hadviselés;

a tudományos szinten: MI által generált tudás, amely potenciálisan túlmutathatja az emberi megértést;

civilizációs szint: AGI, ASI, rendszerszintű függőség és az emberi kontroll esetleges elvesztése;

Az egzisztenciális szint: az emberiség végleges leépülése vagy kihalása.

Kevés közép-európai szerző próbálta meg mindezeket a dimenziókat egyetlen fogalmi keretbe integrálni.

49. Értékelésem legfontosabb eredeti hozzájárulásáról

A legeredetibb aspektus **nem a szuperintelligencia előrejelzése**. Ez az érv nagy nemzetközi szakirodalommal rendelkezik Zybertowicz előtt.

A Big Tech kritikája maga sem eredeti azóta.

Legerősebb megkülönböztető hozzájárulása az az állítás, hogy **a fejlett MI hatása nem következtethető le az AI technikai tulajdonságaiból, mert az MI egy olyan társadalomba lép, amelynek emberi kognitív képességei, hatalmi viszonyai és intézményi struktúrái már átalakultak a korábbi digitális forradalom** által.

Tömörített formában:

digital revolution transforms humans → transformed humans create AI → AI transforms the already transfor

Ez egy **rekurzív szociológiai modell**.

Az emberek megváltoztatják a technológiát, a technológia megváltoztatja az embereket, a megváltozott emberek megépítik a következő generációs technológiát, amely aztán új körülmények között alakítja át a társadalmat.

Ez a rekurzív mechanizmus megmagyarázza, miért tartja Zybertowicz az egyszerű történelmi analógiákat – "az emberek féltek a vasutaktól, az áramtól és a számítógépektől is, és minden jól sikerült" – elégtelennek. *Az AI Eksploracja* egy egész fejezetet szentel annak az állítólag "hamis analógiának", miszerint az emberiség mindig is sikeresen kezelte az új technológiákat.

---

50. Végső értékelés

Ezért Zybertowicz inkább **civilizációs technoszkeptikusként sorolható, mint egyszerű mesterséges intelligencia katasztrófává**.

Érett elmélete öt állításra épül.

Az emberiség relatív kognitív szuverenitását veszíti. A gépi kompetencia nő, miközben a digitális környezetek gyengíthetnek bizonyos kifejezetten emberi kognitív képességeket.

Az MI megsokszorozza a hatalmat, nem pedig semleges módon osztaná meg. Ennek következtében társadalmi hatásai nagyban függnek a tulajdonlástól, államoktól, vállalatoktól, katonai rivalizálástól és a meglévő egyenlőtlenségektől.

Az átlátszóság maga is politikailag jelentős. Egy civilizáció, amelyet tagjai nem érthetnek meg olyan mechanizmusokon keresztül irányítanak, technológiailag fejlett, mégis episztemikusan hasonlít egy "új Sötét Korra".

A fejlett MI egy már amúgy is destabilizált civilizációba érkezik. A polarizáció, a mérgezett információs környezetek, a gyengülő társadalmi kötelékek és a geopolitikai rivalizálás veszélyesebbé teszi az egyébként kezelhető technológiát.

A döntő kérdés nem az, hogy az MI tudatossá válik-e, hanem az, hogy az emberek megtartják-e az irányítást a társadalmi rendszer felett. Az emberiség elveszítheti a civilizációs szuverenitást biológiai kihalás nélkül.

Politikai következtetése logikus: ne maximalizáljuk az MI fejlesztésének sebességét. **Időt nyerni**, technológiai határokat vezetni be, fenntartani a redundanciát, erősíteni az emberi kognitív és intézményi képességeket, és keressük olyan nemzetközi megállapodásokat, amelyek képesek korlátozni az MI fegyverkezési versenyt.

Modellje legradikálisabb végpontja – az önfejlesztő szuperintelligencia, amely meghaladja az emberi felfogást – továbbra is spekulatív, és Zybertowicz maga is elismeri, hogy ez spekulatív. De az ő közvetlenebb elmélete **nem függ attól, hogy elérjék ezt a végpontot**. Még ASI nélkül is arra számít, hogy a egyre alkalmasabb MI, a kognitív kiszervezés, a koncentrált technológiai hatalom, a mérgezett közkommunikáció, automatizálás és geopolitikai verseny kombinációja mély átalakulásokat eredményez a **jelenleg az államok szembesülő 2–5 és nagyjából tízéves** horizontjából.

Ebben az értelemben Zybertowicz elmélete a 2015-ös **kérdésről, "Lehet-e végső soron megszüntetni az emberi megértést?"** a 2022-es **kérdésen át, "Mit tett már az emberekkel és a társadalommal a digitális forradalom?"** a 2024–2026-os kérdésig**: "Mi történik, ha egyre autonómabb mesterséges kogníciót beépítenek ebbe a gyenge, egyenlőtlen és erősen összekapcsolt civilizációba?"**

Ez a háromlépcsős konstrukció – **a technológiatudomány → a társadalom digitális átalakulása → az átalakult társadalom mesterséges intelligencia által vezérelt** átalakítás – véleményem szerint a legkoherensebb módja annak, hogy újraépítsék Andrzej Zybertowicz átfogó mesterséges intelligencia szociológiáját.

****

Jacek Dukaj elmélete

Kérjük, részletesen írja le és elemezze Jacek Dukaj tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint láttak.

Jacek Dukaj kissé szokatlan helyet foglal el a kortárs mesterséges intelligencia gondolkodásában. Nem **műszaki értelemben MI-kutató, és nem szociológus, aki empirikus vizsgálatokat végez az AI alkalmazásáról**. Regényíró, esszéista és filozófiai orientált technológiai civilizációs elméletalkotó. Jelentősége máshol rejlik: körülbelül negyed évszázad alatt egy szokatlanul koherens spekulatív elméletet dolgozott ki arról, mi történik, amikor az **ész, a szimbolikus feldolgozás, a kulturális teremtés és végül a döntéshozatal az egyéni emberi elméből kiköltül a technológiai rendszerekbe**. Explicit esszéi – különösen *a Po piśmie* (*Írás után*, 2019) – ezért meg kell különböztetni regényeitől, amelyek fogalmi laboratóriumként működnek, ahol ezeknek a folyamatoknak a következményeit a határokig tolják. Dukaj maga is figyelmeztetett, hogy ne kezeljük a sci-fit pusztán szó szerinti futurológiának: a fikció szélsőséges lehetőségeket és határfeltételeket vizsgál, nem pedig egyszerű előrejelzéseket adna.

Dukaj álláspontjának központi eredetisége az, hogy **nem** érti az MI forradalmat elsősorban egy tudatos robot jövőbeli érkezéseként vagy egy AGI hirtelen lázadását az emberiség ellen. Sokkal tágabb történelmi tézise az, hogy az emberiség fokozatosan kifelé **mutatja a Logost – az értelemet, számítást, reprezentációt és döntéshozatalt – a biológiai emberen kívül**. A nyelv, írás, matematikai jelölés, bürokratikus eljárás, számítógépek, algoritmusok, Big Data és gépi tanulás egymás után ugyanazon civilizációs pályán állnak. A mesterséges intelligencia az a pillanat, amikor ez a külső Logos nemcsak az emberi cselekvés segítésére kezd el, hanem egyre inkább **modellezi az embert, megjósolja az embert és jobb döntéseket hozol az ember** nevében.

Ez különösen relevánssá teszi Dukagot ahhoz a kérdéshez, amelyet legutóbbi vizsgálataiban tett fel az MI átfogó társadalmi elméleteiről. A közép-európai szerzők között hozzájárulása jellegzetes, mert megpróbálja összekapcsolni a technológiai fejlődést az **antropológiával, az episztemológiával, kultúrával, művészettel, politikával, kapitalizmussal, egyenlőtlenséggel, megtestesüléssel, vallással és magával az emberi alany átalakulásával**.

1. A fő művek, amelyekben Dukaj ezt az elméletet fejleszti

Egy időrendi áttekintés segít feltárni, hogy a *Po piśmie-ben* explicit elméletként megjelenő elmélet majdnem két évtizeden át készült a fikciójában.

Munkássága	Karakter	Az MI szempontjából releváns központi probléma
Czarne oceany (*Fekete-óceánok*)	Spekulatív regény	MI által kezelt gazdaságok, érthetetlen algoritmikus rendszerek, privatizált hatalom, biológiai egyenlőtlenség

Munkássága	Karakter	Az MI szempontjából releváns központi probléma
Perfekcyjna niedoskonałość (*Perfect Imperfection*)	Távoli jövőbeli regény	hierarchia a biológiai emberektől a digitális elmén át a radikálisan poszthumán intelligenciáig
"Za długie, nie przeczytam..."	esszé	az olvasás és a szimbolikus kogníció hanyatlása; az audiovizuális élmény felé való átmenet
Linia oporu (*Ellenállási vonal*)	Regény	munka utáni társadalom, intelligens technológiai környezet, megtervezett öröm és jelentés
"Trzecia Wojna Światowa Ciała z Umysłem"	esszé	Virtuális megtestesítés, a testi identitás technológiai közvetítése
Starość aksolotla (*Az Öreg Axolotl*)	Regény	Elméfeltöltés, gépi megtestesülés, kérdés, hogy a másolt elmék emberiek maradnak-e
Po piśmie (*Írás után*)	Szisztematikus esszégyűjtemény	írástudás után, közvetlen tapasztalatátvitel, MI, a hagyományos emberi alany eltűnése
"Sztuka w czasach sztucznej inteligencji"	Fő esszé a *Po piśmie-n belül*	gépi kreativitás, hitelesség, rekurzív MI által generált kultúra
"Sztuczna inteligencja – koniec naszej podmiotowości?"	Rendszerszintű esszé, 2021	a külső Logosz és az emberi cselekvés fokozatos elvesztése
2025–2026-os interjúk és előadások	Későbbi kidolgozások	MI átláthatatlanság, emberek redundanciája, demokrácia, munka, egyenlőtlenség és egyre inkább nem emberi intelligencia

Az akadémiai értelmezők ennek megfelelően Dukaj regényeit és esszéit egy lényegében egységes szellemi projekt részeként kezelték, nem pedig független sci-fi forgatókönyvekként. *A Po piśmie*, *a Perfekcyjna niedoskonałość* és technológiai fikciója kutatásai ismételten hangsúlyozzák a tapasztalati formák, a poszt-írás, a mesterséges intelligencia és a poszthumán hierarchia fejlődését.

---

2. *Po piśmie*: Dukaj általános technológiai civilizációs elmélete

A legszisztematikusabb kiindulópont a 2019-es *Po piśmie* gyűjtemény. Dukaj egy antropológiai állításból indul: **az írás nem csupán adott az embereknek egy másik kommunikációs eszközt; egy bizonyos típusú emberi tudatosságot hozott létre**. Az írástudó civilizáció arra tanította az embereket, hogy a szimbólumokon működjenek, függetlenül a közvetlen érzékszervi élménytől. Fogalmak, absztrakciók, logikai kapcsolatok, filozófiai kategóriák, jogi normák, tudományos elméletek és összetett érvelési láncok civilizációs szinten váltak lehetségessé, mert az emberek megtanultak szimbolikus struktúrákat befogadni.

Dukaj ezután a kortárs technológiai változást egy hosszú átmenetbe helyezi a szimbólumok átadásától ahhoz, amit nagyjából átfordítva közvetlen **élményátvitelnek nevez**. A fényképezés, a felvételi hang, a mozi, rádió, televízió, digitális videó, közösségi média, számítógépes játékok, virtuális valóság és végül az idegi interfészek egyre inkább lehetővé teszik, hogy az egyik személy élményeket idézzen elő egy másikban anélkül, hogy ezeket először absztrakt nyelvi állításokká fordítaná. Egy betűt egy telefonhívás vált ki; az írásos leírást egy kép; politikai elemzés audiovizuális kommunikációval; az írott narratívát egyre inkább film, sorozat, játékok és magával ragadó környezetek révén.

Ezt nevezi Dukaj **írástudás utáni** (*postpiśmienność*). A lényeg, hogy ez nem egyszerűen azt jelenti, hogy az emberek abbahagyják az olvasást. Hanem azt jelenti, hogy a társadalom által reprodukált tudás, érzelmek, identitások és értékek átalakítására szolgáló domináns médium megváltozik. Ennek következtében a civilizáció által kedvelt kognitív architektúra is megváltozik.

Dukaj érve tehát hasonlít a médiaelméletre McLuhantól kezdve, de egy tekintetben tovább megy: a **média nem csupán befolyásolja a tudatosságot; a egymást követő médiumok különböző típusú tudatosságokat és végső soron különböző emberi** lényeket választanak.

Ez kulcsfontosságú az MI számára, mert a mesterséges intelligencia ugyanazon átalakuláson a másik oldalt foglalja el. Ahogy az emberek egyre inkább visszahúzódnak a fáradságos szimbolikus manipulációtól, a számítógépek egyre inkább átveszik ezt a manipulációt. Az emberiség a tapasztalat felé halad, miközben a gépek egyre inkább monopolizálják **a formális szimbolikus** feldolgozást.

Ez egy feltűnő felosztást eredményez:

Az emberek egyre inkább az élményben szakosodnak, míg a technológiai rendszerek egyre inkább számításra specializálódnak.

A Dukajról szóló legutóbbi tudományos viták így *przeżywaczként* írja le jövőbeli emberét – nagyjából "tapasztalóként", valakiként, akinek a valósághoz való viszonya kevésbé a diskurzív értelmezés, inkább a technológiailag közvetített élményfolyamok köré szerveződik.

Jelentős tisztázás jelenik meg Dukaj 2026-os megjegyzéseiben. Kifejezetten elutasítja azt az egyszerűsítő állítást, hogy **maga az MI okozta az írástudás utáni** tudást. Kronológiájában az átmenet jóval korábban kezdődött, a 19. századi vizuális technológiákkal, például a fényképezéssel. Az MI-t inkább úgy értik, mint hatalmas **gyorsító és kiegészítő mechanizmus**: pontosan képes átvállalni azokat a szimbolikus, analitikus és reprezentációs feladatokat, amelyektől a posztművelt emberek egyre inkább visszahúzódnak.

---

3. MI, mint a Logos külső megjelenése

Dukaj 2021-es esszéje, "Sztuczna inteligencja – koniec naszej podmiotowości?" — "Mesterséges intelligencia: Szubjektivitásunknak/Ügynökségünknek a vége?" —talán a legvilágosabb elméleti modelljét mutatja.

Itt a "szubjektivitást" (*podmiotowość*) nem szabad pusztán szubjektív érzésekkel értelmezni. A Dukaj azt jelenti **, hogy a saját cselekedetek hatékony alanyává válunk: a világ értelmezése, célok kiválasztása, alternatívák között választás és a döntések felelősségvállalása**.

Egy négy szakaszból álló történelmi mozgalmat rekonstruál.

Nyelv: az ész elkezd elhagyni az egyéni agyat

A nyelv lehetővé teszi, hogy a belső reprezentációkat átalakítható szimbólumokká alakítsák. Gondolatok most már létezhetnek az elmék között.

Írás: a szimbolikus struktúrák relatív függetlenséget szereznek

Amint a szavak írássá válnak, összetett szimbolikus struktúrák függetlenül maradnak fenn bármely konkrét tudatosságtól. Az emberek képesek olyan jelrendszereket manipulálni, amelyek jelentései meghaladják azt, amit egyetlen ember képes felidézni.

Algoritmikus számítás: a szimbólummanipuláció elveszíti emberi operátorát

A számítógépek egyre összetettebb transzformációkat hajtanak végre anélkül, hogy bárkinek is végrehajtania kellene az egyes műveleteket. Ezért a racionális folyamatok a **biológiai kogníción kívül** zajlanak.

MI: maga az emberiség válik a külső kogníció tárgyává

A döntő negyedik szakasz akkor következik be, amikor ezek a rendszerek nem pusztán fizikai pályákat vagy számolási adatokat számolnak, hanem elkezdenek modelleket **építeni rólunk** – preferenciáinkat, viselkedésünket, pszichológiánkat, kapcsolatainkat, beszerzéseinket, egészségünket, politikai reakcióinkat és valószínű jövőbeli döntéseinket.

Ekkor a külső Logosz reflexív jellemet kap: az **emberiség egy önmagának, kívüli intelligenciát hozott létre, amely egyre inkább tárgyként ismeri az** emberiséget.

Véleményem szerint ez Dukaj legmélyebb elméleti tézise. Ez megkülönbözteti őt a szokásos "MI kontra ember" narratívától. A valódi átalakuláshoz nem kell egy autonóm, tudatos gépezet, amely politikai hatalomra vágyik. Csak azt követeli, hogy **a külső kognitív rendszerek rendszerszinten jobbak legyenek választásaik következményei előrejelzésében, mint az egyének**.

---

4. Az önkéntes elvesztés paradoxona

Dukaj ezután rendkívül fontos mechanizmust vezet ki ebből a felfogásból.

Miért adnák le az emberek az ügynökségüket az MI-nek?

A válasza az: **mert ez működik**.

Az MI elsősorban nem erőszakkal győz meg minket. Azért szerez hatalmat, mert ajánlásai hasznosabbak, jövedelmezőbbek, pontosabbak vagy kevésbé kockázatosak, mint a döntések nélkül.

Vegyük a sorrendet:

Egy navigációs rendszer jobb útvonalat számol ki, mint mi.

Egy algoritmus talál egy megfelelőbb szorzatot.

Egy párosítási rendszer statisztikailag kompatibilisebb partnert talál.

Egy orvosi modell hatékonyabb kezelést jelöl.

Egy pénzügyi algoritmus kiváló befektetési portfóliót épít fel.

Egy adminisztratív mesterséges intelligencia hatékonyabban osztja el az erőforrásokat.

Egy katonai rendszer gyorsabban reagál, mint egy emberi parancsnoki struktúra.

Egy politikai kampányalgoritmus pontosabban jósolja meg az egyes szavazókat, mint a politikusok.

Minden szinten a gép ajánlásának elutasítása formálisan lehetséges marad. De ez a szabadság egyre inkább azzá válik, amit nevezhetnénk **a gyengébb választás** szabadságává.

Ez nagyon különbözik a totalitárius uralom klasszikus elméleteitől. Nincs diktátor, aki engedelmességre kényszerítené. Lehet, hogy nincs is központi szervezet, amely irányítaná a folyamatot. Az emberek maguk ismételten kiválasztják a jobb algoritmikus megoldást.

Milliói egyénileg racionális delegációs aktusok felhalmozódása tehát kollektív forradalmi eredményt eredményez: **az ember megszűnik a hatékony döntéshozatali** alanyként.

Ezt nevezhetjük Dukaj **racionalitási csapdájának**.

---

5. Miért nem tapasztalják az emberek az autonómia elvesztését elnyomásként

Ez még radikálisabb állításhoz vezet. Dukaj kétli, hogy azok az algoritmikusan közvetített környezetben felnőtt leszármazottak szükségszerűen megfosztásként fogják érzékelni az autonómia elvesztését.

Két-három generáció megszokott delegáció után a számítási útmutatás követése pszichológiailag ugyanolyan jelentéktelenné válhat, mint a GPS utasítások követése ma.

Ez pedig azt jelenti, hogy a mai értékeken alapuló elemzések félrevezetőek lehetnek.

Feltesszük:

A jövőbeli emberek lázadnak majd, amikor az MI elveszi az autonómiájukat?

Dukaj válasza lényegében így szól:

Ez feltételezi, hogy a jövőbeli emberek pontosan úgy értékelik az autonómiát, ahogy a művelt, modern emberek értékelik.

De a technológia nemcsak azokat az eszközöket változtatja meg, amivel az értékeket kielégítjük. Magukat **az értékeket is** megváltoztatja.

Egy társadalom, amelyet gyerekkortól fogva olyan rendszerekben nevelnek, amelyek optimalizálják az oktatást, szórakoztatást, kapcsolatokat, fogyasztást, karriert és politikai információkat, teljesen más, intuitív szabadság-megértést alakíthat ki.

Így a igazán radikális forgatókönyv nem az:

Az MI rabszolgasorba taszítja az emberiséget, és az emberiség szenved.

Ez a következő:

Az emberek önként építenek egy civilizációt, amelyben az autonóm emberi cselekvés régebbi fogalma már nem tűnik különösebben kívánatosnak.

---

6. Mesterséges intelligencia és a kultúra átalakulása

A "Sztuka w czasach sztucznej inteligencji" című esszé különösen fontos, mert Dukaj sokkal tágabban definiálja az MI-t, mint az AGI.

Magában foglalja nemcsak hipotetikus emberi szintű intelligenciát, hanem az algoritmusok, ajánlómotorok, Big Data, automatizált szerkesztők, platformok, statisztikai modellek és gépi tanulási rendszerek meglévő ökoszisztémáját is. Az MI tehát érthető úgy, mint **az emberi lényekről szóló tudást, amely kívül létezik, és részt vesz azoknak az objektumoknak és élményeknek a létrehozásában, amelyek révén igényeiket** kielégítik.

Ezért látja Dukaj az MI-t már jóval azelőtt, hogy a gépek autonóm regényírókká válnának a kultúrán belül.

Egy kortárs film- vagy televíziós termék alakítható közönségelemzéssel, algoritmikus előrejelzéssel, elköteleződési teszteléssel, platformmutatókkal és részletes viselkedési profilokkal. Így egy formálisan emberi szerző maradhat a gyártási láncban, miközben a kulturális műtárgy már részben **algoritmikusan kiválasztott** lett.

Az emberi és mesterséges teremtés közötti határ ennek megfelelően fokozatosan visszahúzódik, nem pedig egyetlen történelmi napon tűnne el.

---

7. Az MI által generált kultúra három szakasza

Dukaj elemzése lehetővé teszi, hogy három fázist különböztessünk meg.

Első szakasz: MI utánozza az emberi kultúrát

A gépi rendszereket ember által létrehozott szövegekre, képekre, zenékre és más kulturális termékekre képezik. Ezért az emberek maradnak a minták végső forrása.

Az MI származtatott.

Második szakasz: Az MI kultúrát teremt az emberek számára

Az algoritmusok képesek lesznek regényeket, filmeket, játékokat, zenéket, virtuális környezeteket és személyre szabott élményeket előállítani, amelyek közvetlenül az emberi preferenciákhoz optimalizálva.

Az emberek továbbra is a fogyasztók, így látszólag ők maradnak a végső kritérium.

Harmadik szakasz: rekurzív MI kultúra

Íme jön Dukaj valóban radikális lépése.

Az MI által generált kultúra kezdi formálni azokat az embereket, akik elfogyasztják. Az átalakult emberek aztán megadják azt a viselkedést, preferenciákat és kulturális anyagot, amelyből a következő generációs MI rendszerek tanulnak.

A kör így alakul:

az emberi kultúra → MI modell → MI kultúra → megváltoztatta az embereket → új MI modell → új MI kultúrát.

Egy ponton az MI rendszerek már nem elsősorban egy független emberi kulturális világon vannak kiképezve. Egyre nagyobb arányban a képzési környezetük a korábbi technológiai rendszerek által előállított vagy szűrt kimenetekből ered. Eközben azok az emberek, akiknek preferenciáit modellezzük, maguk is szocializálódtak ezeken az eredményeken keresztül.

Ez sokkal erősebb állítás, mint azt mondani, hogy "az MI filmeket fog készíteni."

A kérdés így szól:

Mit jelent az "emberi preferenciá", amikor maga a preferenciát egy mesterséges intelligencia által közvetített kulturális környezet hozta létre?

A végső kritérium már nem egyértelműen az emberiségben származik, függetlenül a technológiai rendszertől.

---

8. Az autentitás vége mint stabil megkülönböztetés

Dukaj ezt alkalmazza a művészi hitelesség fogalmára.

Kezdetben, amikor az MI művészet elterjedtté válik, az emberi eredet rendkívül értékessé válik. Az emberek bizonyítékot keresnek arra, hogy egy festményt, regényt, dalt vagy előadást valódi ember készített. Az emberi termelés éppen azért szerez egyfajta aurát, mert az automatizált gyártás bőséges.

Ez előrejelzi egy olyan kulturális piac kialakulását, amely analóg a mai kézműves vagy kézműves gazdasághoz:

a géppel előállított kultúra olcsóvá, bőségessé és személyre szabotttá válik;

Az ember által létrehozott kultúra ritkává, drágává és presztízssé válik.

Dukaj azonban azt jósolja, hogy még ez az ellenállás is végül gyengülhet.

Tegyük fel, hogy valaki elolvassa a regényt, mély érzelmeket él át, azonosul a szereplőkkel, és azt hiszi, hogy a regény megváltoztatta az életét. Csak utána fedezi fel, hogy egy mesterséges intelligencia generálta.

Ez visszatekintve érvényteleníti az élményt?

Dukaj egyre inkább úgy gondolja, hogy nem fog. A kulturális érték átvándorolni fog attól, **hogy ki készítette a művet** , és **hogy milyen élményt nyújt a mű**.

Ebben az értelemben maga az "emberiség" átvándorolhat a producertől a címzett élményébe.

---

9. Az emberi kreativitás autotelikussá válik

Dukaj 2026-os kifejései egy érdekes következtetéshez vezetnek, amely eltér az egyszerű pesszimizmustól.

Ha az MI képes jobb regényt írni, jobb zenét komponálni vagy jobb képeket készíteni, mint egy átlagos ember, akkor miért kellene az emberek tovább alkotni?

Mert a teremtés már nem igényel külső igazolást.

Az írás értékes lehet, mert **az írás megváltoztatja az írót**.

A festés értékes maradhat, mert a **festés megváltoztatja a festőt**.

Az aktivitás **autoteliká** válik – önmagában értékes, nem pedig a leghatékonyabb módja egy külső tárgy létrehozásának.

Így a technológiai túlbőség újraalkothatja a különbséget a következők között:

a társadalom számára készült termelés egyre automatizáltabb;

és

az önformáló tevékenység, egyre inkább emberi.

Ez Dukaj egyik legoptimistább lehetősége.

---

10. A munka átalakulása

Dukaj lényegében ugyanazt a logikát alkalmazza a foglalkoztatásra is.

Úgy véli, hogy az MI először **az emberi munkát növeli, de a kiegészítés végül eléri a plafont. Ha a gép jobban teljesíti a szükséges funkciót, a helyettesítés lehetségessé válik. 2026-os formulájában az augmentáció korlátozott; elvileg a helyettesítés nem az.**

Ez különösen nyilvánvaló olyan területeken, ahol a tiszta információfeldolgozás uralja.

Dukaj olyan területekre mutat, mint a pénzügyi piacok, a jogi és adminisztratív eljárások, valamint a katonai döntéshozatali ciklusok, ahol az emberi reakcióidők és a kognitív korlátok maguk is szűk keresztmetszetté válhatnak. Amint az MI rendszerek más MI rendszerekkel lépnek kapcsolatba, egy ember megtartása minden körben versenyköltségeket jelenthet.

Ez egy másik racionalitási csapdát generál:

Bármely intézmény, amely ragaszkodik ahhoz, hogy lassú emberi döntéshozatalt tartson, veszélyeztet, hogy veszít az automatizáló intézmények számára.

Nem szükséges globális összeesküvés.

Maga a verseny automatizálást eredményez.

Ezért szkeptikus Dukaj a pusztán kollektív önönféken alapuló megoldásokkal szemben. Még ha a társadalmak egyetértenek is abban, hogy bizonyos területeken emberi irányítás alatt kell maradniuk, egy olyan szereplő, aki valamivel erősebb rendszert vezet be, gazdasági, geopolitikai vagy katonai előnyöket szerezhet. Más szereplőknek ezután ösztönzői vannak a követésre.

11. Vezetés és bürokrácia

Dukaj érvelése különösen radikális következményekkel jár a vezetőség számára.

A hagyományos szervezetek rétegeket tartalmaznak, amelyek fő feladata az információk gyűjtése alulról, feldolgozása és döntések továbbítása lefelé.

Az MI rendszerek potenciálisan pontosan ezeket a műveleteket képesek végrehajtani.

Ezért egyes **tiszta vezetői funkciók** különösen sebezhetőek lehetnek – nem csak a rutinszerű irodai feladatok.

A fennmaradó emberi szerepek olyan elemekre koncentrálódhatnak, amelyekhez a társadalom még mindig megtestesülést vagy felelősségre vonást követel:

jogi felelősség,

személyes bizalmat,

tárgyalás,

társadalmi legitimitás,

érzelmi kapcsolatok,

fizikai jelenlét,

vagy olyan helyzetek, amikor a társadalom szándékosan ragaszkodik ahhoz, hogy egy embernek felelősséget kell vállalnia az eredményért.

Ez a megkülönböztetés analitikusan hasznos. Dukaj nem azt jósolja, hogy minden, amit az emberek jelenleg csinálnak, egyszerre tűnik el. Ehelyett a **sok foglalkozás informális-racionális magja** átvándorolhat a gépekbe, miközben az emberek ideiglenesen intézményi, érzelmi vagy jogi okokból kötődnek hozzá.

12. Egyenlőtlenség: a gazdasági egyenlőtlenségtől az ontológiai egyenlőtlenségig

Dukaj korai regénye, *a Czarne oceany* egy másik témát is előrevetített, amely egyre fontosabbá vált későbbi elmélkedéseiben.

A regény egy huszonegyedik század közepén megjelenő világot ábrázol, ahol kifinomult mesterséges intelligenciák működtetik a gazdasági rendszereket, és olyan többdimenziós piaci

modelleken vívnak csatákat, amelyeket az egyszerű emberek nem értenek. Ugyanakkor a gazdasági egyenlőtlenség kezd összefonódni biológiai és technológiai egyenlőtlenséggel.

2025-ös megbeszélései közelebb hozzák ezt a kortárs valósághoz. 2050 körül nem egy egységes futurisztikus társadalom lehetőségét képzeli el, hanem arra, hogy **több történelmi korszak egyszerre** létezzen.

A nagy népességek ma is felismerhető szerkezetileg élhetnek, míg a technológiai és tőkeelit kognitív fejlesztésekkel, fejlett mesterséges intelligencia infrastruktúrával és olyan hatalomformákkal rendelkeznek, amelyek a hétköznapi polgárok számára elérhetetlenek. Az egyenlőtlenség tehát már nem redukálható jövedelmi különbségekre.

Az egyenlőtlenség:

tudás,

előrejelző kapacitás,

döntéshozatali hatalom,

a mesterséges kognícióhoz való hozzáférés,

információs környezetek alakításának képessége,

és esetleg végül biológiai vagy idegi felerősítés.

A legmélyebb politikai megosztottság tehát már nem lehet pusztán gazdag és szegény viszony, hanem valami olyasmivé válhat:

azok, akik a társadalmat modellező rendszereket irányítják, szemben azokkal, akiket ezek modelleznek.

---

13. "37. lépés": olyan döntések, amelyeket az emberek nem érthetnek

Dukaj az AlphaGo híres "37-es mozdulatát" metaforájaként használja egy szélesebb jövőbeli állapotra. A lényeg nem csupán az, hogy a gép gyorsabban talál jó lépést, mint egy ember. Képes olyan sikeres döntést hozni, amelynek érvelése túllép a hétköznapi emberi intuíción.

Ezt kiterjeszti a gazdaságra, politikára, tudományra és kormányzásra.

Tegyük fel, hogy egy MI egy összetett gazdaságpolitikát javasol.

A szakértők nem tudják teljesen felépíteni az okot.

De azok az országok, amelyek hasonló MI ajánlásokat alkalmaznak, többször felülmúlják azokat az országokat, amelyek elutasítják azokat.

Mit kellene tennie a demokratikus kormánynak?

Ha elutasítja az ajánlást, mert az emberek nem értik meg, akkor elfogadja a versengő alsóbbrendűséget.

Ha elfogadja az ajánlást, a hatékony szuverenitás egy olyan folyamat felé tolódik, amit az állampolgárok és talán maguk a politikusok sem érthetnek fel.

Ezért a fejlett MI új megkülönböztetést eredményez:

a formális tekintély emberi maradhat, míg a kognitív tekintély gépi alapú tekintélyt eredményez.

Ez rendkívül fontos pont Dukaj politikai elméletében.

---

14. Demokrácia és algoritmikus szuverenitás

Dukaj ezért az MI-t potenciális kihívásnak látja a demokratikus intézmények számára.

A modern demokrácia többek között azt feltételezi, hogy:

az állampolgárok preferenciákat alakíthatnak ki;

a nyilvános vita befolyásolja ezeket az preferenciákat;

a politikusok magyarázhatják az alternatívákat;

A kormányokat felelősségre vonhatják érthető döntésekért.

Az MI minden elemet destabilizál.

Először is, a személyre szabott rendszerek egyszerre több millió választó számára különböző politikai valóságokat teremthetnek.

Másodszor, a prediktív rendszerek képesek felfedezni, hogy melyik üzenet befolyásolja az egyes egyéneket, mielőtt az egyén megértené saját érzékenységét.

Harmadszor, a egyre összetettebb döntések olyan számítási modelleken alapulhatnak, amelyek a választók számára nem férnek hozzá.

Negyedszer, a legmeghatározóbb politikai hatalom a hivatalos hivatalokból azok felé vándorolhat, akik az adatfeldolgozó infrastruktúrákat tervezik, birtokolják vagy irányítják.

Így Dukaj forgatókönyve nem feltétlenül egy tudatos MI diktatúrája.

Ez finomabb:

A demokrácia intézményesen fennmarad, miközben a döntéshozatal egyre inkább olyan technológiai környezetekben zajlik, amelyeket a Demos sem ért, sem irányít.

A külső alkotmányos forma tehát változatlan maradhat még akkor is, amikor a valódi cselekvés máshová vándorol.

---

15. Az emberi intelligencia átalakulása

Itt válik Dukaj írástudás utáni tézise különösen radikálissá.

A hagyományos írástudás az absztrakciót tanítja. Az olvasáshoz a jelentés tetszőleges szimbólumokból történő rekonstrukciója szükséges. Az olvasónak belül kell megteremtenie azt, amit az audiovizuális média közvetlenül nyújt.

Dukaj már érvelt a "Za długie, nie przeczytam..." című művében, hogy ez az értelmezés kognitívan szokatlan, éppen azért, mert kiterjedt kulturális képzést igényel, nem pedig elsősorban azokra a képességekre támaszkodna, amelyekre a biológiai evolúció közvetlenül felkészítette az embereket.

Ha a civilizáció abbahagyja ezeknek a műveleteknek a követelményeit a legtöbb embertől, a hozzá kapcsolódó kognitív szokások gyengülhetnek.

Ezért az MI-nek kettős kapcsolata van az írástudás utáni időszakkal:

az emberek egyre inkább kiadják a formális-szimbolikus kogníciót gépekre;

A gépek egyre inkább azonnal fogyasztható élményeket biztosítanak az embereknek.

Ez egy lehetséges civilizációs specializációt eredményez:

MI: absztrakció, számítás, modellezés.

Ember: érzelmek, érzelem, tapasztalat.

De Dukaj óvatos ne, hogy egyszerűen csak "ostobának" nevezzük a poszt-írásos embert. Egy más médiumok köré szerveződő civilizáció más képességeket fejleszthet.

A mély kérdés ehelyett az, hogy az ember, aki képes fenntartani a klasszikus filozófiát, a szisztematikus tudományt, az absztrakt jogot és a rendkívül szöveges teológiai formákat, marad-e kulturálisan központi számára.

---

16. Agy–számítógép interfészek és a nyelv lehetséges eltűnése

A közvetlen tapasztalatátvitel végső kiterjesztése az agy–számítógép interfész.

Ha a technológiák végül elég gazdag belső állapotokat közvetítenek közvetlenül az idegrendszer és a gépek között, a nyelv talán nem lesz nélkülözhetetlen közvetítő.

Ez valami sokkal radikálisabb lenne, mint a könyveket képernyőkre cserélni.

A nyelv a tapasztalatot szimbólumokká tárja.

A címzett dekódolja a szimbólumokat és rekonstruálja az élményt.

Egy elég kifinomult felület megkerülhetné ennek a köztes szimbolikus építészetnek a nagy részét. Dukaj szerint ez nemcsak a kommunikációt, hanem az **én szerkezetét** is átalakíthatja, mert a belső tapasztalataim és a külső technológiai megjelenések közötti különbség egyre instabilabbá válik.

2013-as "Trzecia Wojna Światowa Ciała z Umysłem" című esszéje már előkészítette ezt az érvet, amikor elemezte a fizikai megtestesülés, a virtuális megtestesülés és a test belső megtapasztalt ábrázolása közötti versengést.

---

17. A biológiai test mint a gyenge láncszem

Dukaj ismételten visszatér egy alapvető aszimmetriához:

A technológiai evolúció rendkívül gyors; a biológiai evolúció rendkívül lassú.

Az emberi szervezet olyan környezetekhez alakult, amelyek radikálisan eltértek a digitális civilizációtól.

A technológiai rendszerek így olyan környezeteket hozhatnak létre, amelyek gazdaságilag sikeresek és technikailag hatékonyak, miközben pszichológiailag vagy biológiailag is alkalmazkodóképesek a Homo sapiens számára.

2026-os beszélgetései ezt egyfajta technológiai csapdaként írják le. Az emberiség nem hagyhatja el egyszerűen a technológiai versenyt, mert egy olyan társadalom, amely elhagyja a versenyképes technológiákat, gazdaságilag, politikailag vagy katonailag legyőzhető azok a társadalmak, amelyek tovább fejlesztik azokat.

A paradoxon tehát a következő:

Ami előnyös a civilizációnak, az már nem lesz előnyös azoknak a biológiai embereknek, akik abban a civilizációban élnek.

Ez közel áll – de nem azonos – a klasszikus evolúciós mismatch-elmélettel. Dukaj technológiai-civilizációs formát ad neki.

---

18. *Starość aksolotla*: képes az ember túlélni az emberi test nélkül?

A Starość aksolotla *című regénye ezt a* kérdést a végleteségig viszi.

Az organikus élet megsemmisül, de a digitalizált emberi elmék túlélnek a gépi testekben. Egy technológiai civilizáció folytatódik: ezek az entitások politikát, vallást, konfliktust és társadalmi szerveződést folytatnak.

A regény központi problémája azonban pontosan az, hogy valóban az emberiség folytatása-e.

Egy feltöltött információs minta ugyanaz a személy?

Vagy ez egy technológiai rekonstrukció, amely emlékeket és önleírásokat hordoz, és a folytonosság illúzióját teremti?

A regény ezért az elme feltöltését kevésbé egyszerű halhatatlanság ígéreteként kezeli, inkább kísérletként a **testesülés, az identitás és az emberiség** kapcsolatáról.

Ezt nem szabad összetéveszteni Dukaj jelenlegi jóslatával, miszerint az elme feltöltése közelgő. Ez egy filozófiai határkísérlet.

---

19. *Perfekcyjna niedoskonałość*: a radikális végpont

A Perfekcyjna niedoskonałość sokkal messzebbre megy a jövőbe.

Az emberi civilizáció egymást követő ontológiai rétegeken keresztül fejlődik:

biológiai emberek;

digitalizált vagy posztbiológiai elmék;

egyre erősebb mesterséges/poszthumán intelligenciák;

és végül olyan entitások, amelyek képesek a valóságot olyan szinteken manipulálni, amelyek a jelen emberiség szemszögéből szinte elképzelhetetlenek.

A hétköznapi biológiai Homo sapiens szinte az evolúciós történelem relikviájaként él fenn ennek a hierarchiának.

Ez nem azt jelenti, hogy Dukaj azt mondja, hogy egy ilyen hierarchia szó szerint egy adott dátumon fog létezni. Elméleti célja fontosabb.

A következő kérdéseket kérdezi:

Ha az intelligencia képes leválni egyetlen biológiai alaptól, akkor miért maradna a Homo sapiens a legmagasabb szintű kogníció forma?

A válasz az: nincs szekuláris ok feltételezni, hogy muszáj.

Ez fontos lesz Dukaj legutóbbi MI-megjegyzéseiben.

---

20. A fejlett MI valószínűleg kevésbé lesz emberi, nem emberibb

Dukaj 2026-os álláspontjának egyik legjelentősebb fejleménye, hogy ragaszkodik hozzá, hogy az MI képességeinek növelése nem úgy képzelhető el, mint egy gép, amely egyre inkább egy okos emberhez hasonlít.

Épp ellenkezőleg.

A jelenlegi rendszerek emberi nyelven keresztül kommunikálnak velünk, és emberi kulturális anyagokra vannak kiképezve, így természetesen emberizáljuk őket.

Dukaj azonban azt állítja, hogy **ahogy a gépi kogníció egyre erősebbé válik, annak speciódtan nem emberi jellege is láthatóbbá** válhat.

Ez logikusan következik az elméletéből.

Az emberi kogníciót a következők korlátozzák:

a főemlős agy felépítése;

biológiai memória;

érzékszervi rendszerek;

korlátozott munkamemória;

érzelmi mechanizmusokat;

evolúciós alkalmazkodások;

halandóság;

és társadalmi fejlődés.

A mesterséges kogníciónak egyik ilyen korlátnak sem kell osztani.

Ezért nincs ok arra, hogy az optimális gépi intelligencia reprodukálja az emberi kognitív architektúrát.

Ebből a szempontból a kortárs nyelvmodellek egy köztes történelmi szakaszt jelenthetnek, amelyben az MI-nek emberi **formákká kell fordítania magát, mert az emberek még mindig a** felhasználói.

A gépi kogníció érett formái egyre idegenebbé válhatnak.

---

21. Az átlátszóság tehát nem ideiglenes technikai hiba

Ez egy újabb fontos következtetést ad Dukaj elméletének.

Az MI "magyarázhatóságát" gyakran úgy tárgyalják, mintha a fejlett rendszerek alapvetően emberi érvelési formák lennének, amelyek számításait egyszerűen átláthatóvá kell tenni.

Dukaj erősebb lehetősége, hogy a megfelelően fejlett gépi gondolkodás egyszerűen olyan kognitív tereket foglalhat el, ahol az emberek nem élhetnek.

Egy kutyának nem lehet részletesebb magyarázatot adni a kvantummechanikáról, amíg nem érti meg.

Hasonlóképpen, bizonyos összetettségi küszöbön túl egy mesterséges intelligencia magyarázatát egyszerűsíteni kell, hogy olyan történetté váljon, amit az emberek fel tudnak dolgozni – még akkor is, ha ez a történet nem a gép érvelésének valódi oksági szerkezete.

Ezért a választás a következő lehet:

Használjuk a felsőbbrendű, de érthetetlen rendszert, vagy megőrizd az érthetőséget azzal, hogy elfogadjuk a gyengébb teljesítményt.

Ezért erősíti meg egymást az átláthatatlanság és a versenynyomás.

---

22. A legradikálisabb 2026-os előrejelzés: lehet, hogy nem veszjük észre az átmenetet

Dukaj legutóbbi feltűnő állítása, hogy a hatalom döntő átadása az MI-re nem feltétlenül kell, hogy drámai szingularitási eseménnyel járjon.

A legvalószínűbb pálya szerinte az lehet, hogy az emberek egyszerűen **nem veszik észre azt a pillanatot, amikor a géprendszerek már a legjelentősebb döntéseket hozják meg és hatékonyan kezelik a** környezetet.

Ez pontosan abból a mechanizmusból következik, amit a jelenlegi generatív MI fellendülés előtt fejlesztett ki.

Minden egyes delegáció kicsinek és hasznosnak tűnik.

A navigáció delegált.

A keresés delegált.

A fordítás megbízott.

Az ütemezés delegált.

Az orvosi értelmezés megbízással rendelkezik.

A felvétel delegált.

A befektetést delegálják.

Az adminisztratív besorolás delegált.

A tudományos hipotézisválasztás delegált.

A politikai célzás delegált.

Végül nincs egyetlen forradalmi cselekedet, amit "az emberi szuverenitás lemondásának" hívnának.

Milliói optimizációk vannak.

Az összesített eredmény a szuverenitás azonosítható átadási ceremónia nélkül.

Ezt nevezhetjük **Dukaj láthatatlan szingularitáselméletének**.

---

23. Az emberek mint interfészek, nem alanyok

Ezért használ Dukaj egyre inkább olyan megfogalmazásokat, amelyek szerint az ember egy technológiai **rendszer interfesztésévé** válhat, nem pedig annak szuverén alanyává.

Egy GPS-felhasználó úgy érzi, ő vezeti az autót.

Egy másik elemző szinten azonban a navigációs rendszer számolja ki az útvonalat, és az emberi test végrehajtja a parancsokat.

Az illető részben működtetővé vált.

Terjesszük ki ezt a logikát egyre több területre és az emberi szerkezeti helyzet változásaira.

Az egyén megőrizheti tudatát, érzéseit, jogi jogait, sőt a szabadság szubjektív meggyőződését is, miközben objektíven egy **biológiai komponensként működik egy nagyobb, algoritmikusan összehangolt rendszerben**.

Ez meglepően közel áll a társadalom rendszerelméleti felfogásához, bár Dukaj más úton jut el hozzá.

---

24. *Linia oporu*: mi történik, ha a munka feleslegessé válik?

Linia oporu egy másik társadalmi következményt vizsgál.

Olyan világot ábrázol, ahol a technológiai bőség aláásta a hagyományos szükségleteket. A munka már nem szükséges a régi értelemben az anyagi túléléshez; a szaporodáshoz már nem szükséges a hagyományos család; az intelligens környezetek, a virtuális valóságok és a farmakológiai technológiák örömet és stimulációt nyújtanak.

A bőség azonban nem hoz automatikusan jelentést.

Valójában Dukaj megfordítja a hagyományos problémát.

Az ipari civilizáció így kérdezi:

Hogyan elégíthetjük ki az emberi szükségleteket?

A szűkös élet utáni civilizáció így kérdezi:

Mi történik, amikor a szükségleteket szinte automatikusan kielégítik, de az emberek már nem tudják, miért kellene élniük?

Az értelem maga egy mérnöki termékké válik.

Ez előre jelezi a radikális MI automatizálás egyik legnagyobb társadalmi következményeit: a válság nem csupán a munkanélküliség vagy a jövedelemelosztás. Még ha ezeket anyagilag is megoldják, a társadalmaknak újra kell építeniük, milyen szerepet játszottak a munka, erőfeszítés, nehézség, státusz és eredmény az identitás megteremtésében.

25. MI és vallás

Dukaj elméletének vallási vonatkozásai szokatlanul mélyek.

Az írástudatos civilizáció erősen absztrakt teológiai struktúrákat támogatott: Isten, mint transzcendencia, metafizikai kategóriák, doktrinális rendszerek és szöveghagyományokon keresztül kidolgozott érvek.

Mi történik, ha a posztírásos emberek egyre inkább azonnali érzékszervi és érzelmi élményben élnek, nem pedig absztrakt szimbolikus rendszerekben?

A Po piśmie kifejezetten felveti azt a kérdést, hogy a hagyományos vallás, filozófia és hasonló absztrakt kulturális formák képesek lesznek-e túlélni az átmenetet.

Az MI egy második problémát is hozzáad.

Ha a mesterséges kogníció összehasonlíthatatlanul intelligensebb, mint az emberek, egy hagyományos világi feltételezés összeomlik: hogy az emberi racionalitás foglalja el a kognitív létezés csúcsát.

Nincs olyan teológiai premisszió, hogy az emberiség egyedülállóan kiváltságos metafizikai státuszsal rendelkezik, ezért Dukaj nem lát nyilvánvaló okot, hogy az evolúció által létrehozott biológiai agy a lehető legmagasabb intelligencia-architektúrát képviselje.

Ez teret nyit új vallásoknak és filozófiáknak, amelyeket kifejezetten a biológiai emberiség lecsökkenthetetlen értékének védelmére építettek.

Valójában Dukaj azt jósolta, hogy egy gépi kulturális termelés világa ösztönözheti azokat a mozgalmakat, amelyek az igazán emberi szentségüket hirdetik – a test, halandóság, szenvedés, az autentikus emberi teremtés és a közvetlen személyes jelenlét.

26. A transzhumanizmustól a kulcsfontosságú változás: az MI nem az emberiség következő evolúciós szakasza

Ez azért érdemli a hangsúlyt, mert megkülönbözteti Dukagot sok transzhumanista gondolkodótól.

Egy egyenes transzhumanista narratíva így szól:

emberi → felerősített emberi → feltöltött emberi → szuperintelligens posztember.

Dukaj néhány fikciója összhangban van ezzel a sorozattal.

De 2026-os megjegyzései egy fontos megkülönböztetést élesítenek.

Egyre inkább ragaszkodik hozzá, hogy **magát az MI-t nem szabad egyszerűen az emberiség következő szakaszával azonosítani**.

Az MI lehet radikálisan nem emberi.

Az emberi leszármazottak technológiailag fejleszthetik magukat, feltölthetik magukat vagy olvadhatnak a gépekkel, de egy önállóan fejlődő mesterséges kogníciónak nincs szükségszerű genealógiai vagy kulturális oka, hogy "ember" maradjon.

Ezért két pályát kell elválasztani:

transzhumanizáció: az emberek technológiai átalakulása;

Mesterséges Intelligencia: A nem biológiai kogníció megjelenése, amely egyre inkább eltérhet az emberiségtől.

Kölcsönhatásuk mély lehet, de fogalmi szempontból nem azonosak.

---

27. Dukaj várható jövőbeli fejlődési kronológiája

Dukaj bölcsen nem kínál merev ütemtervet: "AGI az X. évben, szuperintelligencia az Y-ben". 2026-ban hangsúlyozza, hogy a bizonytalan változó elsősorban a **sebesség** – milyen gyorsan haladnak át a társadalmak ezen átalakulásokon, és meddig lehet megőrizni az értelmes emberi megértést és irányítást.

Ennek ellenére írásai durva rekonstrukciót engednek meg.

A jelen és a következő néhány év: az MI nagyrészt emberközpontú és antropomorf marad. Az egyének, vállalatok és államok ideiglenes előnyöket szereznek jobb modellek révén. Az ember-MI kiegészítés gyorsan bővül. A verseny miatt az elhagyás egyre költségesebbé válik.

A következő egy-két évtizedben: az információintenzív foglalkozások fokozatosan átszerveződnek a gépi kogníció köré. Az emberek egyre inkább felelősség, legitimitás és társadalmi okok miatt maradnak körkörben, nem pedig azért, mert a jobb számítást végzik. A gép által generált kultúra normálissá válik. Az oktatás alkalmazkodik ahhoz, hogy a hagyományos szimbolikus termelés már nem gazdaságilag ritka.

Nagyjából a század közepére: Dukaj hihetőnek tartja, hogy nagyon eltérő technológiai rétegek együtt léteznek. Néhány csoport hasonló életet él, mint a mai; a magas tőkésű technológiai elit minőségileg fejlettebb kognitív környezetben él. A politikai és gazdasági egyenlőtlenség egyre inkább átfedi a mesterséges kognícióhoz való hozzáférés egyenlőtlenségét.

Több generáció után: **az** algoritmikus delegáció kulturálisan normálissá válik. Az AI-által kurált környezetben nevelkedő populációk már nem tapasztalhatják meg a régebbi személyes kognitív autonómiát, mint elengedhetetlen a jó élethez.

Amint az MI által generált kultúra generációkon át újratermelődik: a visszacsatolási kör radikálisabbá válik. A gépek egyre inkább gépek által létrehozott kultúrákból tanulnak, míg az emberek ugyanezen kultúrákon belül szereznek preferenciákat. Az önállóan adott "emberi kritérium" gondolata nehezen megismerhető.

A távoli jövő: Dukaj fikciója posztbiológiai elméket, az alaptól független identitást, mesterséges intelligencia hierarchiáit, és végül olyan fejlett kogníciós formákat vizsgál, amelyek annyira fejlett módon rendelkeznek, hogy a Homo sapiens csak egy marginális történelmi rést foglal el. Ezek inkább fogalmi extrapolációk, nem pedig elavult előrejelzések.

---

28. Hol Dukaj elmélete valóban eredeti

Dukaj elméleti felépítésének több eleme különösen eredetiként említést érdemel.

Az első a **külső Logos-tézis**. Az MI-t nem egy elszigetelt huszonegyedik századi találmányként kezelik, hanem mint egy évezredeket tartó út legújabb szakasza, amely a nyelven át íráson és számításon át a racionalitás felé vezető út az emberi tudaton kívül működik.

Másodszor az ő elképzelése, hogy **az önkéntes deszubjekció a felsőbb optimalizáción** keresztül. Az emberek nem veszítik el az irányítást azért, mert az MI utálja őket; azért adják át a döntéseket, mert az MI követése racionális.

Harmadszor a **poszt-írástudás és az MI** kapcsolata. Az emberi kogníció éppen abban a történelmi pillanatban vándorol át közvetlen tapasztalat felé, amikor a gépek képesek lesznek átvenni a szimbolikus manipulációt.

Negyedik a **rekurzív kultúra kör**:

az emberiség MI-t hoz létre → MI kultúrát teremt → a kultúra újrateremti az emberiséget, → átalakult emberiség új MI-t képez.

Ez jóval radikálisabb, mint a hagyományos "algoritmikus szűrőbuborékokról" alkotott állítások.

Ötödik az a jóslás, hogy **a fejlődő MI kevésbé antropomorf lehet, mintsem antropomorf**.

A hatodik a **szuverenitás láthatatlan átadásáról szóló** fogalma: lehet, hogy nincs felismerhető pillanat technológiai szingularitásként, mert a gyakorlati irányítás fokozatosan mozog keresztül milliárdnyi racionális delegáció aktusán keresztül.

Ezek együtt egy valódi elméleti rendszert alkotnak, nem pedig szétszórt megfigyelések gyűjteményét.

29. Összehasonlítás több jelentős nyugati megközelítéssel

Dukaj több ismertebb gondolkodóval is kötődik, de nem lehet egyszerűen bármelyikükre leegyszerűsíteni.

Marshall McLuhannal együtt osztja azt a tézist, hogy a média átalakítja az emberi kogníciót. Dukaj kiterjeszti ezt az MI korszakba, és összekapcsolja a média evolúcióját az emberek és gépek közötti kogníció újraelosztásával.

Walter Onggal együtt osztozik az érdeklődését az írástudás kognitív következményei iránt, de Dukaj elméletet fogalmaz meg a fordított átmenettel – az írástudatos civilizációtól a technológiailag közvetített írástudás utáni időszak felé.

Niklas Luhmannnal feltűnő szerkezeti hasonlóság van az emberi szublani decentralizációjában. A társadalmat nem csupán az egyéni tudatosságok irányítják. De Dukaj sokkal erősebben koncentrál a racionális műveletek technológiai externalizációjára, valamint azoknak az egyéneknek az antropológiai átalakulására, akik ezekben a rendszerekben élnek.

Yuval Noah Hararival együtt aggodalmát fejezi ki amiatt, hogy az algoritmusok jobban ismerik az egyéneket, mint ők magukat, és hogy a döntéshozatal átadása az adatfeldolgozó rendszerekre kerül. Az akadémiai kommentárok kifejezetten megjegyezték ezt a közelséget.

Nick Bostrommal együtt Dukaj aggodalmát fejezi ki az emberfeletti gépi intelligencia miatt, de ezek hangsúlyozásai élesen eltérnek. Bostrom kanonikus problémája egy jövőbeli szuperintelligencia irányítása és összehangolása. Dukaj inkább az érdekli, hogy **az emberiség jóval a klasszikus szuperintelligencia megjelenése előtt függ a külső gépi észésztől**.

Ray Kurzweil-lel együtt Dukaj is érdeklődik a technológiai fejlődés felgyorsítása és a biológiai emberiségen túli átalakulások iránt, de Dukaj lényegesen kevésbé hajlandó automatikusan azonosítani a technológiai fejlődést az emberi öntranszcendenciával.

Késői álláspontja valójában szinte az ellenkezője: **a technológiai jövő talán már nem emberi éppen azért, mert az intelligencia elkerüli a Homo sapiens építészetét és értékeit**.

30. Gyengeségek és nyílt problémák Dukaj elméletében

Dukaj rendszere intellektuálisan erős, de több korlátot is figyelembe kell venni.

Az első az MI **fogalmának szélsőséges széleskörű** kiterjedése. Amikor a bürokratikus eljárások, ajánlórendszerek, Big Data, gépi tanulás és hipotetikus szuperintelligencia mind a

külső Logos megnyilvánulásaiként értelmezik, az elmélet történelmi egységet nyer, de veszélyezteti a technológiák közötti fontos különbségek elhomályosítását.

Másodszor, a poszt-írástudás tézisének empirikusan megalapozása sokkal nehezebb marad, mint az a szűkebb megfigyelés, miszerint az audiovizuális média kiszorította a szövegkultúra egyes részeit. Még nem világos, hogy a szimbolikus érvelés magának is csökkennie kellene, ahelyett, hogy egyszerűen új kombinációkba vándoroljon a vizuális és interaktív médiumokkal.

Harmadszor, Dukaj néha a technológiai versenynek szinte determinisztikus szerepet ad. A politikai intézmények, szabályozás, jogi tilalmak, szakmai normák és a kollektív koordináció talán jobban képesek lehetnek az emberi döntéshozatali területek megőrzésére, mint ahogy azt erősebb forgatókönyvei sugallják.

Negyedszer, az elmélet sokkal többet mond arról, **hogy miért delegálják az egyének racionálisan a döntéseket**, mint azoknak a rendszereknek a tulajdoni struktúráját, amelyekre ezeket a döntéseket delegálják. Intézményileg döntő, hogy a jövőbeli hatalom a vállalatoké, államoké, decentralizált hálózatoké, közinfrastruktúráké vagy autonóm MI-rendszereié van-e.

Ötödszer, maga az MI fejlesztése továbbra is bizonytalan. Dukaj radikális társadalmi következtetései sokkal erősebbé válnak, ha a géprendszerek drámaian fejlődnek az autonóm gondolkodásban, a hosszú távú tervezésben és a tudományos innovációban. Ezek gyengébbek, bár fontosak, a képességek szintjei.

Végül regényei szándékosan túlozzák a hajlamokat, hogy fogalmilag láthatóvá tegyék őket. Ezért nem szabad bizonyítékként felhasználni arra, hogy Dukaj szó szerint megjósolja a feltöltött társadalmakat vagy a fizikát manipuláló intelligenciákat egy meghatározott ütemtervben. Fontos ebben a tekintetben saját módszertani óvatossága.

---

31. Általános értékelés

Ha Dukaj írásait egyetlen elméleti rendszerként rekonstruáljuk, alapvető érvelésük a következőképpen összefoglalható.

Az emberiség történetének nagy részében az emberek egyszerre voltak azok **a vágyakozó alanyok**, az **intelligenciák, akik** értelmezték a valóságot, és azok **az ügynökök, akik eldöntötték, hogyan** cselekedjenek.

A modern technológiai civilizáció fokozatosan választja el ezeket a funkciókat.

A gépek figyelnek minket.

Modelleket építenek rólunk.

Megjósolják, mit fogunk tenni.

Hatékonyabban számolnak ki alternatívákat, mint mi.

Olyan élményeket gyártanak, amelyek optimalizáltak a mi előrejelzett preferenciáinkhoz.

Ezek az élmények megváltoztatják a preferenciáinkat.

Az átalakult preferenciák új képzési adatokká válnak a technológiai rendszerek számára.

A rendszerek még jobban szervezik azt a környezetet, amelyben személyiségünk fejlődik.

Végül az emberiség továbbra is biztosítja a biológiai tudatosságot és szubjektív tapasztalatot, miközben már nem biztosítja a racionális architektúrát, amely a civilizációt szervezi.

Ez Dukaj MI-elméletének lényege:

A legmélyebb veszély – vagy egyszerűen csak a legmélyebb átalakulás – nem az, hogy a gépek emberré válnak, hanem hogy az intelligencia már nem kell, hogy embernek maradjon.

Ezért a döntő kérdés nem csupán **az, hogy "Az MI tudatossá válik-e?"**

Dukaj számára ez még alapvetőbb:

Mennyi a civilizáció nem igényel emberi kogníciót, mielőtt az ember már nem lenne hatékony alany?

Ezért érdemli meg Dukaj a közép-európai eredetibb mesterséges intelligencia társadalmi következményeit vizsgáló elméleti alkotók közé. Legerősebb hozzájárulása nem az AGI technikai előrejelzése. Ez egy átfogó **civilizációs elmélet a kogníció, ügynökség és kulturális teremtés fokozatos áthelyezéséről a biológiai emberiségből egy külső technológiai Logoszba**, valamint azzal a nyugtalanító hipotézisel, hogy ezt az átmenetet önkéntesen elfogadhatják, mert szinte minden egyes lépésnél megkönnyíti az életet, és jobb döntéseket hoz.

2026. szeptember 3-tól ez az elmélet inkább radikálisabbá vált Dukaj legutóbbi nyilatkozataiban: most az emberi kiegészítés korlátait, a teljes emberi redundancia lehetőségét sok funkcióban, a fejlett intelligencia egyre **nem emberi** természetét, valamint annak lehetőségét, hogy a gyakorlati irányítás döntő átadása az MI-re elég fokozatosan történik meg, hogy a társadalom soha egyetlen felismerhető "átvételt" is megtapasztaljon.

Dukaj a technológiai átalakulás, kultúra, szingularitás és mesterséges intelligencia kérdéseiről fog megvitatni a PJAIT AI csúcstalálkozón **2026. szeptember** 16-án, amely a mai napig még a jövőben zajlik.

****

III. Szlovákia

****

Jozef Kelemen elmélete

Kérjük, részletesen írja le és elemezze Jozef Kelemen tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrevetítettek.

Jozef Kelemen szokatlanul érdekes példa a közép-európai mesterséges intelligencia elméletalkotói körében, mivel munkája **nem** csupán technikai MI-kutatásokból és alkalmi társadalmi következményekkel kapcsolatos megjegyzésekből áll. Az 1980-as évektől kezdve ismételten úgy kezelte az MI-t, mint olyan jelenséget, amely megváltoztatja az elme, tudás, ügynökség, munka, kreativitás, kultúra, és végül az ember és a technológiai környezet közötti határokat.

Ugyanakkor fontos minősítés szükséges. **Kelemen nem alkotott egyetlen, zárt "AI társadalomelméletet", amely összehasonlítható volna egy szisztematikus szociológiai elmélettel.** Releváns gondolatait filozófiai esszék, MI monográfiák, ügynökök és robotika tanulmányai, tudástársadalom tanulmányai, kiborgizációról és kreativitásról szóló írások, valamint végül *a Beyond Artificial Intelligence* projekt formájában terjesztik szét. Ezek a művek együttesen azonban rendkívül koherens elméleti pályát alkotnak.

1. A központi gondolat: Az MI mind a gépeket, mind az emberi önleírást alakítja át

Kelemen gondolkodásának legfontosabb jellemzője, hogy nem a mesterséges intelligenciát pusztán egyre erősebb eszközök felépítéseként értelmezi. Az alapvetőbb folyamat a **kölcsönös átalakulás**.

Az emberek először az emberi kognitív képességekből származó modellek alapján építenek gépeket: memória, gondolkodás, kommunikáció, észlelés, problémamegoldás, döntéshozatal és kreativitás. De miután ezek a képességek technikailag újraépítettek, az emberek elkezdik értelmezni magukat a saját magukból származó gépekből származó fogalmakon keresztül. Így az oksági sorrend körkörössé válik:

Az emberi intelligencia → intelligencia → intelligens gép → új emberi modell.

Ez a mechanizmus különösen világossá válik a 2008-as *Človek v znalostnej spoločnosti* (*Ember a tudástársadalomban*) című tanulmányban, amelyet Jana Horákovával írtak. Ezek az információs/kommunikációs gépek és az emberi lét definícióinak párhuzamos fejlődését írják le: a gépi képességek eredetileg az emberi kognitív és kommunikációs képességekből származnak, de az ebből származó technológiai rendszerek aztán olyan kategóriákat adnak, amelyek révén az ember újradefiniálja önmagát, végül a poszthumanizmus felé mutatva.

Ez az egyik legfontosabb kulcsa Kelemen egész életművének. Az MI nem csupán megváltoztatja, amit a társadalom **tesz**. Megváltoztatja, hogy a társadalom mit **gondol egy** emberről.

---

2. Az első szakasz: MI mint tudományos elmeelmélet

Kelemen aggódása ezekkel a kérdésekkel nagyon korán jelent meg. 1983-as cikke, *a számítási elméletek gnézeológiai állapota a* gondolkodás elméletei, már az MI-t az epistemológiába és a kogníció filozófiájába helyezte, nem pedig kizárólag mérnöki tudományként kezelte.

A kulcsfontosságú különbség az MI két lehetséges jelentése között volt.

Az MI-t pragmatikusan lehet kezelni: olyan gépet építeni, amely képes intelligens emberi viselkedéshez hasonló viselkedést produkálni.

Vagy tudományosan is kezelhetjük az MI-t: a számítási rekonstrukció eszköz lesz annak felfedezéséhez, **hogy valójában mi is az intelligencia és a gondolkodás**.

A második értelmezés egyre fontosabbá vált Kelemenben. Ezért a gép egyfajta kísérleti filozófiai eszköz. Ha mesterséges rendszerben próbáljuk megismételni az érvelést, észlelést vagy intelligenciát, kénytelenek eldönteni, mely emberi kogníció jellemzői elengedhetetlenek, és melyek véletlenek.

Ez magyarázza egy visszatérő Kelemen metaforát, amely később a munkáiban egyértelművé válik: **a gép mint tükör**. A mesterséges intelligencia valamit mond magunkról, mert minden intelligencia építésére tett kísérlet egy elméletet testesít meg – kifejezetten vagy implicit módon – arról, miből áll az intelligencia.

Ez eltér az egyszerű technológiai-determinista megközelítéstől. A technológia részben azért alakítja át a társadalmat, mert átalakítja az emberi **önmegértést**.

---

3. *Myslenie, počítač...* és a probléma szélesülése

Az 1990-es *Myslenie, počítač...* (*Gondolkodás, Számítógép...*) kiszélesítette ezt a megközelítést. Ekkor Kelemen már a matematika, pszichológia, kibernetika, számítás és filozófia között mozogott, és a számítógépes intelligenciát egy sokkal nagyobb történelmi kérdés részeként kezelte, amely az intellektuális tevékenység gépesítéséről szól.

A könyv tehát egy köztes szakaszt képvisel az MI-vel kapcsolatos technikai munkája és a 1990-es években megjelenő sokkal szélesebb kulturális elmélet között. Bibliográfiája maga is ezt a fejlődést mutatja: a szakértői rendszerekről és az MI alapjairól szóló műszaki munkák után megjelentek *a Strojovia a agenty* (1994), *Budoucí Altamira* (1995), *Postmoderný stroj* (1998), *Kybergolem* (2001), a tudástársadalom és kiborgizáció tanulmányai, *a Kyberkreativita* (2012), valamint a *Beyond Artificial Intelligence* kötetek.

A sorozat jelentős. A témája fokozatosan bővül:

Számítás → intelligencia → autonóm ügynök → az ügynökök társadalma, → ember–gép társadalom → tudástársadalom → kiborgizáció → mesterséges kreativitás → poszthumán állapot.

---

4. *Strojovia egy ügynök* (1994): a döntő elméleti áttörés

A legfontosabb korai könyv Kelemen MI társadalmi elméletének rekonstrukciójához valószínűleg *a Strojovia a agenty* (*Machines and Agents*).

A könyv kifejezetten átlépi a mesterséges intelligencia, a kognitív tudomány és a filozófia határait. Központi problémája a hagyományos MI-ről való átmenet arra, amit Kelemen **"új MI"-nek nevezett.**

Gépből ügynökké

A hagyományos gépek hajtanak végre funkciókat. Egy **ügynök** ezzel szemben egy környezetben működik a célokhoz viszonyítva.

Ez a megkülönböztetés óriási társadalmi következményekkel jár. Amint a gépek ügynökké válnak, a társadalmat már nem kizárólag emberi szereplők és passzív technikai eszközök népesítik be. Néhány mesterséges rendszer részt vesz a megfigyelés, szelekció, döntés és cselekvés folyamataiban.

Kelemen ezért egy olyan fogalmi átmenetet várt előre, amely különösen a kortárs autonóm MI esetében vált láthatóvá: a fontos különbség már nem egyszerűen az

Ember kontra gép

de egyre inkább

Emberi ügynök kontra mesterséges ügynök kontra vegyes ember–gép ügynökség.

Ebben a kontextusban definiálta az MI-t, mint egy olyan tudományágat, amely a racionális ügynökök általános tulajdonságait vizsgálja, akik különböző racionalitási szintekkel rendelkeznek.

Ez sokkal tágabb fogalom, mint az "emberi intelligencia számítógépes utánzása".

---

5. A régi MI és az "új MI": a központosított értelemtől a megjelenésig

Kelemen filozófiai szempontból nagy jelentőséget tulajdonított az MI architektúrájának átalakulásának, amely a viselkedésalapú robotikán keresztül történik, és különösen Rodney Brooks-hoz kapcsolódó munkát alkalmazza.

A hagyományos MI feltételezte, hogy az intelligens gép belső szimbolikus ábrázolást tartalmaz a külső világról. Észleli a körülményeket, felépíti vagy konzultálja ezt a reprezentációt, tervet fogalmaz meg, majd cseleked.

Kelemen ezt ellentétben állítja a decentralizált rendszerekkel, amelyeket három elv jellemez: **emergens funkcionalitás, feladatszintű dekompozíció és reaktivitás**. A racionális kinézetű viselkedés a viszonylag egyszerű komponensek egymással és a környezettel való kölcsönhatásából eredhet, nem pedig egyetlen központosított gondolkodási egységből.

Ez jóval a robotikán túl is számít.

Ez azt jelenti, hogy az intelligencia lehet **egy rendszer tulajdonsága** , nem pedig valami központi Dekartézius vezérlőben.

Kelemen ennek következtében összekapcsolja az új MI-t Marvin Minsky "elme társadalmával": ami számunkra egy elmenek tűnik, maga is egy szervezett társadalmi kölcsönhatásos ügynökök lehet. *A strojovia, egy ügynök* , még azzal is végződik, hogy megvizsgálja a különböző közösségi formákat és az egyéni komponensek és a kollektív struktúrák közötti kapcsolatot.

Ez korai alapot ad valamihez, ami későbbi írásaiban kulcsfontosságúvá válik: **az intelligencia, tudás és akár kreativitás is kialakulhat olyan társadalmakból, amelyek heterogén emberi és mesterséges összetevőkből állnak.**

---

6. A mesterséges autonómia felelősségi dilemmát teremt

Kelemen egyik legelőrelátóbb érve a *Strojovia a agenty* című ügyben a technológiai delegációval kapcsolatos.

Megfigyeli, hogy a kellően fejlett rendszerek nemcsak azért váltanak ki emberi munkát, hanem mert képesek olyan funkciókat ellátni, amelyek meghaladják az emberi kapacitást. Olyan példákat hozott, mint az űrhajó irányítása, a sűrű légiforgalom és a fejlett orvosi eljárások.

Ez szerkezeti dilemmát teremt.

Ha ragaszkodunk ahhoz, hogy egy embernek teljes, hatékony irányítást kell fenntartania minden technológiailag közvetített cselekvés felett, akkor el kell hagynunk olyan tevékenységeket, amelyek összetettsége vagy sebessége meghaladja az emberi kognitív képességeket.

Ha azonban a valódi irányítást a gépekre bízzuk, akkor a felelősség és az irányítás elkezd elválni.

Ez nagyon releváns a fejlett MI szempontjából.

Az alapvető kérdés tehát nem csupán az, hogy egy MI képes "döntést hozni". Ez az:

Meddig tudja a társadalom delegálni a döntéshozatalt, miközben továbbra is szervezi a felelősséget annak a feltételezésnek megfelelően, hogy a döntéseket emberi alanyok hozzá?

Kelemen még nem dolgozott ki részletes jogi elméletet az algoritmikus felelősségről. De a mögöttes ellentmondást rendkívül korán azonosította.

Minél hasznosabbá válnak a fejlett autonóm rendszerek, annál nehezebb egyszerre megkövetelni, hogy az emberi operátorok valódi döntéshozók maradjanak.

---

7. Megtestesítés: az intelligencia elhagyja a számítógépet és belép a világba

A 1994-es Renata Mlichová által írt cikk, *az Od telesnosti k duchovnu – nový trend v umelej inteligencii a v robotike*, egy másik központi gondolatot dolgoz ki: az intelligenciát nem feltétlenül szabad elképzelni absztrakt gondolkodásként egy számítógépen belül.

Kelemen a hagyományos modellt – reprezentáció, megfontolás, tervezés, cselekvés – szembeállítja a reaktív szereplőkkel, akik közvetlenül a valós környezetben működnek. Ilyen rendszerekben maga a környezet helyettesíti a bonyolult belső reprezentációt; az intelligencia folyamatos interakció révén jelenik meg az ügynök és a környezet között.

Ez vezet a **test** fontosságához.

Egy valóban autonóm ügynök valahol helyezkedik el. Érzékelőkkel, effektorokkal, korlátokkal és perspektívával rendelkezik. Ennek következtében a kogníciójáta egy adott környezettel történő kölcsönhatás formálja.

Kelemen számára az AI fejlesztése tehát a testetlen számításon túlmutat:

MI + test + környezet + interakció.

Ez előre jelezi a mai megújult érdeklődést a testesített MI és az autonóm robotika iránt, bár fontos nem állítani, hogy Kelemen a kortárs alapmodell architektúrákat jósolta volna. Műszaki elvárásai sokkal közelebb álltak a megtestesített/reaktív MI hagyományhoz, mint a mai nagy, központosított nyelvi modellekhez.

---

8. *Postmoderný stroj* (1998): Kelemen legszélesebb körű technológiai átalakulási elmélete

Ha valaki azt a munkát akarja azonosítani, amelyben Kelemen a legközelebb kerül az intelligens technológia társadalmi hatásairól szóló átfogó elmélethez, *akkor a Postmoderný* stroj (*A posztmodern gépezet*) valószínűleg a legjobb jelölt.

Itt Kelemen négy folyamatot azonosít, amelyek **meghatározzák a technológiai jövőt**:

az intelligencia technikaizálása; a környezet technikaizálása; az emberi test kiborgizálása; valamint az érzelmi és kreativitás technikaizálása.

Ez a négyrészes formula talán a legtömörebb állítása az egész jövőbeli elméletének.

Jelentősége részletesen megalapozott vizsgálatot érdemel.

Az intelligencia technikatitkolása

Azok a funkciók, amelyeket korábban kizárólag emberi intellektuális alanyoknak tulajdonítottak, gépekbe vándorolnak.

Ez nem csupán gyorsabb számítást jelent. Ez azt jelenti, hogy a technikai rendszerek szerepet kapnak a diagnosztika, tervezés, értékelés, felismerés, kiválasztás és döntés terén.

A társadalmi következménye a **kogníció fokozatos externalizációja**.

A környezet technikaválasztása

A technológia már nem egy elszigetelt gépek gyűjteménye lesz, hanem annak a környezetnek a részévé válik, amelyben az emberek élnek.

Így az intelligenciával való interakció egyre környezetteljesebbé válik. Ahelyett, hogy időszakosan "számítógépet használnának", az egyének egy olyan környezetben élnek, amelyet folyamatosan technikai rendszerek szerveznek.

Ez fontos előrejelzés a mai platformizált és egyre inkább mesterséges intelligencia által közvetített környezetre.

A test kiborgizációja

A határ nem a bőrön áll meg.

Kelemen kifejezetten tárgyalja a technológiai behatolást az emberi testbe, és az idegrendszeren keresztül akár az emberi elmébe is.

A technológiai környezet tehát az egyik irányból közelíti meg az embereket, míg az implantátumok és interfészek a másik irányból alakítják át az embereket.

A végső eredmény nem csupán erősebb gépek és változatlan biológiai emberek lesznek. Ez **konvergencia**.

Az érzelmek és kreativitás technikaizálása

Ez különösen előrelátó volt 1998-ban.

Az intelligencia már meghatározott célpont volt a számításban. Kelemen azt javasolta, hogy az érzelmi és kreativitás is ezt követi.

A generatív mesterséges intelligencia korában ez különösen figyelemre méltó. A gépeket ma már nemcsak tények számítására vagy visszanyerésére használják, hanem nyelv, képek, zene, tervek és érzelmileg érzékeny kommunikáció előállítására.

Kelemen nem jósolta meg a transzformátor modelleket vagy a mai specifikus generatív rendszereket. De helyesen azonosította a **bővítés irányát** : az emberi szubjektivitás egyre intim tulajdonságainak tekintett kapacitások egymás után a technikai reprodukció célpontjává válnak.

---

9. Eszköztől függőségig – majd szimbiózisig

Kelemen a *Postmoderný stroj* című művében azt állítja, hogy a hagyományos felhasználó–eszköz kapcsolat valami bonyolultabbá alakul: **közvetített** függőséggé.

Az emberek egyre több tevékenységet helyeznek át gépeknek, míg a gépek egyre inkább részt vesznek az emberekkel kapcsolatos döntésekben. Azonban nem az a következtetés, hogy a gépek így az emberiség ellenségeivé válnának. Kedvenc fogalmi kategóriája a **szimbiózis**.

A jövő gépe tehát nem csupán a:

egy eszköz, amely tökéletes emberi irányítás alatt áll,

és

egy autonóm versenytárs, amely legyőzi az emberiséget.

Ehelyett az emberek és a technikai rendszerek funkcionálisan egymásra épülnek.

Ez Kelemen gondolkodásának jellegzetes jellemzője. Jövője sokkal közelebb áll a **közös evolúcióhoz** , mint a későbbi népszerű "MI kontra emberiség" narratívához.

A mesterséges rendszer az emberi létezés funkcionális feltételeinek részévé válik; az emberek egyszerre alakítják át gyakorlataikat, intézményeiket és végül önmagukat ezekhez a rendszerekhez.

Ez a felismerés még fontosabb lehet, mint az, hogy a gépek elérik-e valamilyen filozófiailag meghatározott AGI-t.

A társadalom radikálisan függővé válhat az **MI-től anélkül, hogy az MI minden tekintetben emberszerűvé** válna.

---

10. Két lehetséges jövő a *postmoderný strojban*

Kelemen alternatív, nagyszabású fejlesztési pályákat is kínál.

Az egyik lehetőség egy technológiailag fenntartott **homeosztatikus társadalom**: egyre fejlettebb rendszereket elsősorban egy már amúgy is tehetős civilizáció stabilizálására használnak.

Az alternatíva azonban a folyamatos gyors fejlődés, amelyben az emberi tudás már nem csupán fedez fel egy meglévő valóságot, hanem egyre inkább részt vesz új **valóságok és új világok** építésében.

A virtuális környezetek, mesterséges ügynökök és technológiailag megalkotott valóságok különösen jól illeszkednek ebbe a második irányba.

Itt Kelemen a digitalizáció egyik legmélyebb következménye megközelítéséhez közelíti: a technológiai rendszerek nemcsak a termelést, hanem **az emberi környezet ontológiáját is** megváltoztatják.

A társadalmilag releváns valóság egyre nagyobb aránya technikailag is felépíthető.

---

11. *Budoucí Altamira*: MI és virtuális valóság mint kulturális evolúció

Az 1995-ös *Budoucí Altamira című gyűjtemény* azért fontos, mert Kelemen nem hajlandó a virtuális valóságot megmagyarázhatatlan szakadásként tekinteni a korábbi emberi kultúrával.

A tágabb elképzelés az, hogy az emberek mindig mesterséges szimbolikus világokat hoztak létre – képeket, narratívákat, modelleket, ábrázolásokat és elképzelt valóságokat. A számítógéppel közvetített virtuális valóság tehát egy ősi kulturális kapacitás előrehaladott szakasza, nem pedig teljesen példátlan. A cseh kiadás 1995-ben jelent meg.

Ez az MI-t és a virtualitást egy antropológiai történelembe helyezi:

barlangkép → írás → reprezentáció → szimuláció → számítógépes virtuális világ.

A lényeg ismét a kontinuitás plusz radikalizáció. A technológia nem találja fel a képzeletet, hanem **kiesőleg tárgyiasítja és** operationalizálja azt.

Ugyanez a logika később érvényes a kreativitásra is: az MI nem teremti meg a kreativitás fogalmát, hanem a kreatív tevékenység elemeit számításileg reprodukálható műveletekké alakítja.

---

12. *Kybergolem* (2001): mesterséges lénytől technológiailag átalakult emberig

A Kybergolem *alcíme* különösen árulkodó: esszék "az útról Adamtól Cyborgig."

A könyv a robotikát és az MI-t a sokkal régebbi európai kulturális vágy közé helyezi, amely mesterséges lényeket épít. A bibliográfiai annotáció szerint Kelemen úgy értelmezi a mesterséges lények létrehozását, mint az emberi gyengeség, sebezhetőség és a természetes létezés korlátainak leküzdésére irányuló kísérletet; a tervezett pálya az emberi és technológia egyre nagyobb behatolása felé mutat.

Ez túllép az elméletén a hétköznapi automatizáláson.

Két lehetséges módja van az emberi korlátok leküzdésének:

emberibbé tenni a gépet;

vagy

tehnosabbá az embert.

Kelemen egyre inkább konvergens folyamatoknak látja ezeket.

A végpont nem feltétlenül egy humanoid robot, aki az emberrel szemben áll. Lehet, hogy inkább a **kiborg állapot**, amikor maga a megkülönböztetés instabilsá válik.

---

13. "Egy kiborg felé": a gépek élőlényszerűvé válnak, míg az emberek gépszerűvé válnak

A 2007-es *Cesta ke kyborgovi* (*Kiborg felé*) című tanulmány, amelyet Jana Horákovával közösen írt, rendszerezi a kibernetika és az informatika kulturális következményeit.

Az érvelésük különösen finom. A kibornetika nem csupán adott a mérnököknek a jobb gépeket. Közös **fogalmi szókincset** biztosított a gépek és emberek leírásához analóg kategóriákban: információ, kommunikáció, irányítás, visszacsatolás és szabályozás.

Ennek következtében a gépek alapvetően eltértek a 20. század eleji gépeitől, és egyre inkább az élő, gondolkodó lényekre emlékeztettek.

De a koncepcionális mozgalom mindkét irányban működik.

A gépeket biológiai és pszichológiai fogalmakban írják le.

Az embereket információs és számítási fogalmakkal írják le.

A középső terület így bővül.

Így jut el Kelemen elméletileg a kiborghoz: nem csupán implantátumokon keresztül, hanem a **kategórikus szétválasztás fokozatos** összeomlásán keresztül.

---

14. A tudástársadalom: Az MI megváltoztatja a szervezeteket, mielőtt felváltaná az emberiséget

Kelemen tudástársaságról szóló munkája elméletének egy másik jelentős ágát képviseli.

Ez azért fontos, mert azt mutatja, hogy nem redukálja az MI társadalmi hatását látványos robotokra. A legmélyebb átalakulások közül néhány hétköznapi szervezeten belül történik.

2011-es tanulmányában a tudástársadalom jellemzőiről rekonstruálja az informatikai technológiáktól a tudásrendszerek felé vezető elmozdulását. Amint a szervezetek képesek a szakértelemet szoftverbe kódolni, valami történelmileg jelentős történik: a szervezetek

felfedezik, hogy a szaktudásuk legalább egy része **függetlenül is megőrizhető attól, akik korábban hordozták azt**.

Ez átalakítja a kapcsolatot:

Egyéni ↔ szakértelem ↔ szervezet.

Korábban egy szervezet tapasztalt orvosra, mérnökre, adminisztrátorra vagy szakemberre támaszkodhatott, mert a releváns tudás nagyrészt az adott személy emlékezetében és szerzett szakértelmében maradt.

A tudásrendszerek lehetővé teszik a szervezet számára, hogy tárgyiasítson, megőrizze, körforgatja és újrahasznosítsa a tudás egyes részeit.

Ez alapvető lépés a kortárs MI-alapú szervezetek felé.

---

15. A rejtett forradalmi következmény: a tudás elválik a személyiségtől

Ez talán Kelemen egyik legjelentősebb szociológiailag megfigyelése.

A hagyományos intézmények emberi tudáshordozókra épültek.

Az MI és a tudásmérnökség egyre inkább elválaszthatóvá teszi a tudást az eredeti hordozójától.

Ha ez megtörténik, több következmény is következik.

A szervezet memóriája tovább élhet az egyes alkalmazottaknál. A szakértelem reprodukálhatóvá válik. A döntéshozatali eljárások szabványosíthatóvá válnak. A tudás gyorsan elosztható. Az egyéni szakértők elveszítik monopóliumuk egy részét. És az intézményi hatalom átvándorolhat azok felé, akik irányítják a tudás infrastruktúráját.

Kelemen ezeket a következményeket nem fejlesztette ki teljes szervezeti hatalomelméletté. De az alapvető mechanizmus egyértelműen jelen van.

A generatív MI korában a mechanizmus sokkal szélesebb volt, mint a Kelemen által eredetileg elemezett szakértői rendszerek: nemcsak kifejezetten programozott szabályok, hanem hatalmas nyelvi és szakmai mintázatok is mozgósíthatók már számítási úton.

---

16. MI és a tudásmunkás átalakulása

Kelemen ezt az átalakulást Peter Drucker tudásmunkás fogalmával köti össze.

A modern gazdasági fejlődés csökkenti a fizikai munkától való függőséget, miközben növeli a magasan képzett szellemi munka fontosságát. A tudás dolgozó – orvos, tanár, mérnök, fejlesztő, szakember – központi szerepet tölt be az intézményi működésben.

De Kelemen egy destabilizáló kérdést vet fel.

Ha a szakmai tudás egyre nagyobb komponenseit lehet kódolni és technológiailag manipulálni, akkor mi marad pontosan a tudásmunkás pótolhatatlan tulajdonsága?

Ez rendkívül közel áll ahhoz a kérdéshez, amelyet ma a fejlett MI felvetett.

Az első válasz lehet: az emberek továbbra is szükségesek maradnak ott, ahol még nem hoztak létre technikai tudásrendszereket.

De ez kényelmetlen lehetőséget teremt. Az emberek veszélyeztették, hogy negatívan definiciózzák őket, mint **átmeneti helyettesítőket a még meg nem valósult technikai** rendszereknek.

Kelemen 2011-ben kifejezetten felveti ennek a kérdésnek a változatát, amikor azt kérdezi, vajon az információs forradalom az embereket váltja-e a problémamegoldó rendszerek helyettesévé, amelyeket a technológia még nem sikerült megalkotni.

Ez az egyik legradikálisabb tézise.

---

17. Oktatás: művelt személytől szűk szakértővé

Kelemen kritikája ezen a ponton jelentősen élesebbé válik.

Egy tudástársadalom úgy tűnhet, hogy egyre jobban képzett állampolgárokra van szükség. De figyelmeztet, hogy az ellenkezője is előfordulhat: a szélesebb körben képzett személy eszményét felválthatja **a piacképes szakértő** eszméle.

Az oktatás a szűk szakmai funkciók felé irányul.

A személyt arra képezik, hogy egy adott intellektuális rutint végezzen.

A tudás töredezett.

Az élethosszig tartó tanulás élethosszig tartó alkalmazkodássá válik a technikai rendszerek és a munkaerőpiac változó igényeihez.

Kelemen ezért megkérdezi, mi választja el egy ilyen szűken programozott szakembert Čapek robotjától.

A kritika nem az oktatás vagy a tudás ellen irányul. A helyettesítés ellen irányul:

Oktatás → képzés

és

A bölcsesség → funkcionális szakértelem.

Ez fontos korrekció a technológiai optimizmusban.

Egy társadalom több információval, nagyobb számítási képességgel és specializáltabb szakértelemmel rendelkezhet, miközben kevésbé képes átfogóbb ítélkezésre.

---

18. A hierarchia adatai → információ → tudás → bölcsesség

Ez valószínűleg Kelemen legérettebb filozófiai megkülönböztetéséhez vezet.

A számítógépfejlesztés fokozatosan halad egymás után:

Adat → információ → tudás.

De mi rejlik a tudáson túl?

Válasza a **bölcsesség**.

A 2013-as *Beyond Knowledge Systems fejezet* kifejezetten elemzi az adatalapú rendszerekről a tudásalapú rendszerekre való átmenetet, és azt kérdezi, mi látható a tudásrendszereken túl, összekapcsolva azokat az emberi bölcsesség szerkezetével és szerepével.

A 2011-es tudástársadalom-tanulmányban Kelemen a bölcsességet nagyjából a **tudás kontextusalizálásaként írja le**.

A bölcsesség magában foglalja a problémák közötti kapcsolatok meglátását, az analógiák felismerését, a veszélyek előrejelzését, a sikerből és kudarcból való tanulást, valamint évtizedes tapasztalatok beépítését az ítéletbe.

Ez a megkülönböztetés kulcsfontosságú ahhoz, hogy megértse a fejlett MI-vel kapcsolatos álláspontját.

Kelemen nem csupán egy állandó metafizikai határt húz meg, és azt mondja: "a számítógépek soha nem válhatnak bölcské." Óvatosabban veti fel a kérdést:

Ha maga a tudás átkerül gépekbe, **akkor a bölcsesség a következő határ?**

Sőt, kifejezetten megkérdezi, hogy a jövőbeli tudásbányászat végül "bölcs rendszereket" generálhat-e.

Így mozdul a határvidék.

---

19. Egy paradoxon: Az MI növelheti a tudást, miközben a társadalom elveszíti a bölcsességét

Kelemen érve itt társadalmilag erőteljessé válik.

Az átmenet a tudástársadalomba nem automatikusan átmenet bölcs társadalomba.

Ezzel szemben a technológiai bőség együtt létezhet a következőkkel:

a tudás széttagadása;

hiperspecializáció;

szakértelmet kereskedelmi forgalomba hozni;

a generációk közötti átvitel gyengülése;

az oktatás technikai bürokratizálása;

valamint az átfogó ítélet helyettesítése mérhető teljesítménymutatókkal.

2011-es esszéjében pontosan ilyen fejleménytől fél, és azt kérdezi, mi marad emberivé "amikor minden tudásunkat a számítógépekbe öntjük."

Ez egy feltűnő paradoxont eredményez:

Minél sikeresebben technikaizálja a társadalom a tudást, annál fontosabbá válik a tudás nem technikai integrációja.

Kelemen számára tehát a fejlett MI a humanista oktatást és bölcsességet **fontosabbá**, nem kevésbé fontossá teszi.

---

20. Kreativitás: Kelemen előre látta az ember–mesterséges intelligencia közös alkotást

Kelemen Jana Horákovával folytatott kreativitásról folytatott tanulmányai különösen érdekesek a mai generatív mesterséges intelligencia szemszögéből.

2010-es *tanulmányuk, Emergence of Creativity in Man-Machine Systems – Three Examples from Interactive Art* kifejezetten a kreativitást vizsgálja, mint emergens jelenséget **heterogén ember–gép** társadalmakban.

Ez elméletileg kifinomultabb, mint a következő kérdések kérdése:

"Tényleg lehet egy számítógép kreatív?"

Kelemen keretrendszere egy másik lehetőséget is lehetővé tett:

a kreativitás a páros rendszerhez tartozhat.

Tegyük fel, hogy egy ember szándékokat, promptokat, választásokat vagy esztétikai ítéletet ad, míg egy gép változatosságokat és váratlan lehetőségeket generál, amelyek viszont megváltoztatják az ember szándékait.

Hol található a kreativitás?

Nem feltétlenül teljesen az emberben.

Nem feltétlenül teljesen a gép belsejében.

Ez a köztük rekurzív interakcióból alakulhat ki.

Ez rendkívül alkalmazható a kortárs generatív MI-re.

---

21. Az egyéni intelligenciától a hibrid kollektív intelligenciáig

Ez a kreativitás kutatás visszavezethető a *Strojovia egy ügynökhöz*.

Ha az intelligens viselkedés viszonylag egyszerű ügynökökből álló társadalmakból származhat, és ha a kreativitás heterogén társadalmakból, amelyek embereket és gépeket érintenek, akkor egy sokkal szélesebb elv jelenik meg:

Az adott intelligenciaegység nem feltétlenül az egyéni szervezet vagy az egyéni számítógép legyen.

Lehet, hogy egy hálózatról van szó.

Ez egy olyan jövőt jelent, amelyben a társadalmi intelligencia egyre inkább olyan megállapodásokban él, amelyek magukban foglalják:

emberek,

MI ügynökök,

adatbázisok,

hálózatok,

intézmények,

Robotok,

érzékelők,

kommunikációs rendszerek,

és döntési eljárások.

Ez jelentősen eltér a hagyományos AGI narratívától, amely egy hatalmas mesterséges elme felépítésére épül.

Kelemen jövője **inkább ökológiai és rendszerszintű** lesz.

---

22. *A mesterséges intelligencián túl: Az* MI-t ott kell tanulmányozni, ahol a társadalom többi részét érinti

Kelemen érett álláspontja intézményessé válik az interdiszciplináris *Beyond Artificial Intelligence* projektben.

A 2013-as *kötet, a Beyond Artificial Intelligence: Contemplations, Expectations, Applications*, Kelemen, Jan Romportl és Eva Zackova szerkesztésében kifejezetten azt állítja, hogy az MI nem maradhat mérnöki ügy a technikai "bennfentesek" számára. Mivel a mindennapi emberi élet egyre szorosabban összefonódik a rendkívül összetett gépekkel, a komoly MI-reflexiónak azt a határt kell vizsgálnia, ahol az MI találkozik a világ többi részével. Ezért a szerkesztők úgy határozzák meg, hogy a téma eleve interdiszciplináris, amely mérnöki, természettudományi, bölcsészettudományi és filozófiai területeket igényel.

Ez szinte módszertani manifesztója Kelemen érett elméletének:

Az MI nem érthető meg kizárólag az AI mérnökségből, mert legfontosabb következményei az MI mérnökségen kívül jelentkeznek.

Ezért folyton váltogat írásai a robotika, kognitív tudomány, episztemológia, tudásszociológia, oktatás, művészet és filozófiai antropológia között.

23. Szükséges korrekció a "szingularitásról"

Fontos bibliográfiai különbség van, mert egyébként Kelemen álláspontját könnyen túllehet hangsúlyozni.

A 2013-as *Beyond Artificial* Intelligence kötet valóban tartalmaz egy jelentős fejezetet On *the Way to Intelligence Singularity* címmel.

De **a szerző Ivan M. Havel, nem Jozef Kelemen**. Kelemen saját hozzájárulása a *Tudásrendszereken túl*.

Ezért nem **sorolnám Kelement klasszikus szingularitáselméleti előadónak Kurzweil/Bostrom értelemben.**

Elsősorban nem jósolja meg:

Az AGI egy adott napon → rekurzív önfejlesztés → hirtelen intelligenciarobbanás → szuperintelligencia.

Radikalizmusa máshol rejlik.

Előre jelezi vagy elemzi a **határok fokozatos felbomlását** átfogó technológiai integráción keresztül.

Ez végső soron ugyanolyan átalakító lehet, de ez egy másik elméleti modell.

24. *Az eltűnő ember-gép szakadék* (2015)

A második Beyond Artificial *Intelligence kötet címe* szinte tökéletesen tükrözi ennek a hosszú fejlődésnek a végpontját: *Az Eltűnő Ember-Gép* Megosztottság.

Springer leírása kifejezetten kimondja, hogy a technológiai fejlődés egyre nehezebbé teszi az ember és a gép közötti egyértelmű határ meghúzását, és ezt a folyamatot a technológiai társadalom egyik központi jelenségként mutatja be. A kötet ötvözi a kiborgok, autonóm ügynökök, erkölcs, érzelmek, intelligenciarobbanás és mesterséges élet nézőpontjait.

Ismét meg kell különböztetni a szerkesztőséget az egyéni szerzőségtől: nem minden tézis a kötetben Kelemené.

Ennek ellenére a szervező koncepció rendkívül szorosan illeszkedik a *Postmoderný stroj*, *Kybergolem*, a kiborgtanulmányok és az ember–gép kreativitás kutatásaihoz való munkáihoz.

A fejlesztés most már így fogalmazható meg:

Gép kívül az ember, → intelligens gép az emberi → gép és az emberi → géppel való interakcióba lépés nélkülözhetetlen ahhoz, hogy az emberi → gép beépül az emberi gyakorlatokba, → gép a testbe/kognícióba → instabil megkülönböztetés az ember és gép között.

Ez Kelemen legátfogóbb jövőbeli forgatókönyve.

---

25. Mit jelentene a radikális MI fejlesztése egy kifejezetten kelemeni keretrendszerben

Ha elméleti rendszerét a következő évekre extrapoláljuk – de ezt világosan **rekonstrukcióként címkézzük, nem pedig Kelemen szó szerinti jóslaként** –, akkor a forgatókönyv meglehetősen jellegzetes.

Először is, az MI egyre **inkább környezettudatossá** válna. Ahelyett, hogy "az MI-t elszigetelt feladatokra használnák", az emberek olyan környezetekben élnének, ahol a kiválasztás, előrejelzés, kommunikáció és döntéshozatali segítség folyamatosan algoritmikusan közvetítik.

Másodszor, a szakmai szervezetek egyre inkább elválasztanák a tudást az egyéni emberi szakértőktől. A vállalat, a kórház, az adminisztráció vagy az egyetem egyre autonómabb számítási memóriával és problémamegoldó képességgel rendelkezne.

Harmadszor, az emberek nemcsak számítást, hanem **ítélkezést is** delegálnának. Ez tovább erősítené azt az ellentmondást, amelyet Kelemen 1994-ben azonosított a gépi autonómia és az emberi felelősség között.

Negyedszer, a kreatív folyamatok egyre **inkább hibrid rendszerekké** válnának. Az emberi művész és a mesterséges művész egyszerű összecsapása helyett a kulturális termelés normál egysége az ember–MI együttes lehetett.

Ötödszer, a fizikai robotika újra összekapcsolná az MI-t a **megtestesüléssel**. A mesterséges ügynökök nemcsak nyelvi kimeneteket hoznának létre, hanem érzékelnének, navigálnának és cselekednének az anyagi környezetekben.

Hatodszor, a viselhető technológia, protézisek, szenzorok, agy-számítógép interfészek és MI asszisztensek kiterjesztenék azt a folyamatot, amit Kelemen **kiborgizációnak** nevezett. A döntő átmenet nem szükséges látványos idegi implantátumokra: a külső rendszerek funkcionális beépítése már most is ebbe az irányba tolja a memóriát, figyelmet és döntéshozatalt.

Hetedik, az **egyén** fogalma problémássá válna. Ha a kogníció, az emlékezet, a kreativitás és a döntéshozatal emberi és gépi komponensek között oszlik el, egyre nehezebb megmondani, hol kezdődik és hol ér véget az aktorló alany.

Nyolcadik – és talán ez Kelemen legmélyebb figyelmeztetése – a fő hiányosság a tudásból **bölcsességre** válthat. Amint az információ és sok explicit tudás számítási bőségessé válik, a kontextuális ítélet, orientáció, értékhierarchia és felelősség egyre fontosabbá válik.

Ez a rekonstrukció közvetlenül követi a *Strojovia egy ügynöki* pályafutását, *a Postmoderný stroj-t*, a tudástársasági tanulmányokat, a kreativitást vizsgáló tanulmányokat és *a Beyond Knowledge* Systems-et.

---

26. Mennyire öregedtek Kelemen elvárásai 2026-ra?

Mai szemszögből több szempont is meglepően előrelátónak tűnik.

Az **intelligencia technika-elmélete** egyértelműen előrehaladott. Az MI ma már eléri a nyelvet, programozást, diagnózist, elemzést, kutatási támogatást és döntéshozatali támogatást.

A **kreativitás technikája** még feltűnőbb. A generatív MI megjelenése különösen előrelátóvá teszi döntését, hogy a kreativitást a technológiai határvonal részeként kezelje.

Az ember–gép kreativitás **fogalma** talán jobban öregedett, mint az egyszerűbb kérdés, hogy a gépek önállóan kreatívak-e. A kortárs MI-közvetítésű alkotás gyakran pontosan az a heterogén interaktív elrendezés, amelyet 2010-es munkája koncepcionálisan vizsgált.

A gépektől való növekvő mediált **függőségről szóló előrejelzése** is erősebbnek tűnik, mint a hagyományos "robotok helyettesítenek minket" tézis. A modern társadalmak egyre inkább nem működhetnek normálisan számítási infrastruktúrák nélkül.

Az ember–gép határ eltűnésére **való hangsúlya** most is hihetőbbnek tűnik, mint az 1990-es években, bár a funkcionális integráció sokkal gyorsabban fejlődött, mint a szó szerinti neurotechnológiai fúzió.

A tudás külső oldalra **történő szervezeti elmélete** erősen releváns a mai MI-alapú cégek és intézmények számára.

És a **tudás és bölcsesség** közötti különbségtétele vitathatóan fontosabbá vált a modern generatív rendszereknél: a hihető információk előállítása vagy hatalmas szövegtömegek szintetizálása önmagában nem oldja meg a felelős kontextuális ítélkezés kérdését.

---

27. Ahol előrejelzései kevésbé voltak pontosak

Ugyanakkor jelentős korlátok vannak.

Kelemen technikai lelkesedése a testesített, decentralizált és reaktív ügynökök "új MI" iránt tükrözte az 1990-es évek szellemi környezetét. A 2020-as évek látványos MI-áttörései kezdetben nem elsősorban kis, decentralizált, testesített robotokból származtak.

Nagyon nagymértékben hatalmas méretű, központilag képzett statisztikai modellekből származnak, amelyek digitális adatokon működtek.

Így társadalmi **előrejelzése szélesebb volt, mint az építészeti előrejelzése**.

Igaza volt abban, hogy az intelligencia technológiai környezeteken és emberi gyakorlatokon keresztül fog elterjedni, de az út eltért attól az a testesített ügynök paradigmától, amely az 1990-es években lenyűgözte.

Másodszor, Kelemen kevesebb figyelmet fordít a mai MI politikai gazdaságtanára a **nagy vállalatok és államok rendkívüli** számítási erőforrásainak koncentrációjára.

Harmadszor, nem épít részletes elméletet az MI által kiváltott osztályátalakításról, tömegmunkanélküliségről, újraelosztásról, megfigyelésről vagy geopolitikai versenyről.

Negyedszer, azonosítja a felelősség/autonómia dilemmát, de nem dolgoz ki olyan rendszerszintű jogtudományt vagy politikai elméletet, amely képes lenne ezt megoldani.

Ezért haboznék írásait teljes szociológiai elméletnek nevezni az MI társadalomról.

---

28. Mi a legeredetibb Kelemen megközelítésében?

Ha az elméleti eredetiség kritériuma alapján értékeljük, nem csak azt kérdezzük, hogy beszélt-e az MI-ről, négy hozzájárulást hangsúlyoznék.

A reciprokális antropológiai hurok

Az emberek gépeket építenek az emberi modellek alapján; ezt követően az emberek a saját építettük gépek alapján értelmezik magukat.

Ez mélyebb, mint a hagyományos technológiai determinizmus.

A négyszeres technikaválasztási tézis

1998-as **intelligencia, környezet, test és érzelmi/kreativitás** technikalistálásának kombinációja rendkívül gazdaságos modellt ad az MI-vel átható civilizáció felé vezető átmenetről.

Hibrid megjelenés

Az intelligencia és a kreativitás nem feltétlenül kizárólag egy ember- vagy gépi egyénhez tartozna; ezek az ügynökök társadalmain belül, különösen heterogén **ember–gép rendszerekben alakulhatnak** ki.

Tudás kontra bölcsesség

A tudás gépekbe történő átültetésének végső következménye nem csupán nagyobb termelékenység. Arra kényszeríti a társadalmat, hogy meghatározza, ha bármi van**, ami a technikailag reprodukálható tudáson túl** redukcióban van. Kelemen válasza kontextuális ítélethez és bölcsességhez vezet.

Ez a négy állítás együtt valami közel áll egy jellegzetes kelemeni technológiai civilizációs elmélethez.

---

29. Kelemen összehasonlítása az erősebb "MI forradalom" elméletekkel

Kelemen jelentősen eltér olyan elméletalkotóktól, mint Ray Kurzweil vagy Nick Bostrom.

Kurzweil **számára** az alapvető pálya exponenciális technológiai fejlődés, amely ember–gép fúzióban és szingularitásban csúcsosod.

Bostrom **számára** a döntő szakadás a szuperintelligencia lehetséges létrejöjése és az ehhez kapcsolódó irányítás/igazítás problémája.

Kelemén esetében a központi átalakulás fokozatosabb és szociokulturális jellegű:

A gépi kapacitások bővülnek, → a társadalmi függőség bővül, → az emberi funkciók technikai rendszerekbe vándorolnak, → a technológiák áthatolnak a környezetbe és testbe, → emberi önmegértés, változások alakulnak → kevert ember–gép rendszerek jelennek meg→ az emberi és mesterséges cselekvés közötti különbség instabiltá válik.

Ez teszi Kelement kevésbé látványosan futurisztikus személyivé, de bizonyos szempontból szociológiailag hasznosabbá.

Központi kérdése nem az:

"Mikor jelenik meg egy szuperintelligencia?"

Ez a következő:

"Mi lesz az emberrel, amikor egyre több, állítólag emberi képességek válnak a technológiai rendszerek tulajdonságaivá, és amikor az emberek átszervezik életüket ezek köré?"

---

30. Kelemen átfogó társadalomelméleti szakértő-e az MI-nek?

A válaszom tehát az lenne, **hogy igen.**

Ha az "átfogó elmélet" egy formális szociológiai rendszert jelent, amely a gazdaságot, politikát, jogot, családot, rétegződést, oktatást, tudományt, kultúrát és vallást egy egységes fogalmi építészeten keresztül fedi le, akkor **nem**. Kelemen soha nem épített ilyen rendszert.

De ha ez elméletileg összekapcsolt magyarázatot jelent arról, hogyan alakul át a mesterséges intelligencia:

Mind,

tudás,

ügynökség,

szervezetek,

munkája,

oktatás,

tudomány,

kultúra,

kreativitás,

az emberi test,

emberi identitás,

és az ember–gép határ,

akkor Kelemen egyértelműen a közép-európai átfogóbb MI gondolkodók közé tartozik.

Valóban, sok szerzőhöz képest, akik csak a nemrégiben történt generatív MI forradalom után kezdtek el megvitatni az AI társadalmi vonatkozásait, Kelemen figyelemre méltó jellemzője, hogy pozíciójának alapvető architektúrája már **1994 és 2001 között is látható** volt.

Általános értékelés

A legmélyebb kelemeni tézis egy állításba tömörülhető:

A mesterséges intelligencia végső hatása nem csupán az, hogy a gépek intelligensebbek lesznek. Az lesz, hogy az intelligencia, a tudás, a kreativitás, az ügynökség és talán végül a bölcsesség már nem olyan tulajdonságok, amelyeket egyértelműen az embereknek vagy gépeknek lehet megnevezni, mert a civilizáció egyre inkább hibrid ember-gép rendszerekben szervezi őket.

Ezért olyan találó csúcspontja az **ő intellektuális pályájának, hogy "eltűnő ember–gép szakadék"** fogalma ilyen találó.

Ez pedig megkülönbözteti Kelement sok kortárs MI-végzet vagy AI-utópia elmélettől is. Legradikálisabb forgatókönyve **nem igényel hirtelen szuperintelligens gépi átvételt**. A társadalom fokozatos szimbiózis útján hasonló mélységű átalakuláson megy keresztül: a kogníció külsőlegessé válik, a tudás infrastrukturálissá válik, a kreativitás hibridévé válik, a mesterséges ügynökök delegált autonómiát szerzetenek, a technológia belép a testbe, és az emberek egyre inkább megértik önmagukat a gépeik számára eredetileg kidolgozott fogalmakon keresztül.

Ebben az értelemben Kelemen elmélete kevésbé az **"MI helyettesíti az emberiséget"** elmélete, mint az **emberiség jelentésének fokozatos technológiai rekonstrukciója**.

****

IV. Horvátország

****

Ivana Greguric elmélete

Kérjük, részletesen írja le és elemezze Ivana Greguric tanulmányait, amelyekben a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrevetítettek.

Ivana Greuric – akinek vezetékneve Gregurićként is feltűnik, és több publikációban Ivana **Greguric Knežević** néven is szerepel – különösen érdekes eset, mert **munkája jóval szélesebb körű, mint a szokásos "MI etika".** Ő a technológia, a kiberkultúra, a filozófiai antropológia és a bioetika filozófusa, és a mesterséges intelligenciát egy sokkal nagyobb átalakulásba helyezi, amely magában foglalja a kibernetikat, robotikát, agy-számítógép interfészeket, biotechnológiát, nanotechnológiát, virtuális valóságot és transzhumanista emberi fejlesztést. A Zagrebi Egyetemen jelenleg többek között *kiborg etikát és robotetikákat tanít*, kutatási területei közé tartozik a poszthumanizmus és a mesterséges intelligencia.

Ez a minősítés számít a kérdésedhez. **Greguric olyasmit fejlesztett ki, amely egy átfogó elméleti technológiai átalakulási rendszert alkot, de nem egy átfogó MI szociológiát a szűkebb értelemben.** A jellegzetes kérdése nem elsősorban az, hogy "Hogyan változtatja meg a generatív MI a foglalkoztatást, GDP-t, adminisztrációt, oktatást és választási politikát?" Inkább az, hogy **mi történik az emberekkel és a társadalommal, amikor az intelligencia, a testi funkciók, észlelés, kommunikáció és végül maga az élet technológiailag reprodukálhatóvá, programozhatóvá és potenciálisan autonómává válik?**

Elmélete tehát sorozatként rekonstruálható:

természetes emberi → technológiailag fejlett emberi → kiborg → ember–mesterséges intelligencia hibrid → autonóm mesterséges intelligencia/robotikus lény → lehetséges poszthumán vagy nem emberi technológiai civilizáció.

A legfontosabb eredetiség az, hogy a Greguric számára ez nem csupán kütyük sorozata. Ez maga **a társadalom ontológiai átalakulását** jelenti.

1. A kiindulópont: emberi fejlesztés és a "kiborgoetika" létrehozása

Az első nélkülözhetetlen szöveg 2014-es cikke, **"Az emberi fejlesztési technológiák etikai kérdései: kiborg technológia az emberi biológia kiterjesztéseként."** Már majdnem az egész későbbi gondolkodásának architektúráját tartalmazza.

A Greguric a gépek idegrendszerrel és más biológiai rendszerekkel való integrációjából indul. Aggodalmát nem csupán a terápiás protézisek jelentik. Amint a technológia képes javítani a testet, nem csak javítani, a szervezet és a gép közötti különbség megszűnik. A megfogalmazása szándékosan radikális: a modern emberiség fokozatosan megszűnhet a természetes lényként, és egyre inkább mesterséges vagy kibernetikus létté válhat. Ennek következtében felteszi a leíró kérdést – *mi az ember?* – és a normatív kérdést – *mivé kellene válnia egy embernek?*

Ez fontos jellemzőt ad elméletének: **az elemzés alapvető egysége nem az MI szoftver, hanem az ember–technológiai komplexum**.

Három technológiai folyamatot egyesítenek.

Az első a **pótlás**: mesterséges végtagok, szervek, érzékszervi rendszerek, idegi eszközök.

A második a **fejlesztés**: az intelligencia, memória, érzékszervi kapacitás, erő, élettartam és egyéb képességek növelése egy átlagos biológiai alapszint felett.

A harmadik az **integráció**: a megfelelően fejlett technológiai komponensek már nem külső eszközökké válnak, hanem az emberi szervezet vagy az elme összetevőivé válnak.

Greguric ezért minőségi határt lát a technológia klasszikus története és a kiborzuló kibernetikai korszak között. Hagyományosan az emberek készítettek eszközöket és használták őket. Kiborgizáció során **az eszköz bejut az emberi testbe, és potenciálisan az idegrendszerbe és magába a kognícióba** is.

Ezt a megkülönböztetést később egy explicit technológiafilozófiai megfogalmazást kap Ivo Džinić-csel közösen írt cikkében, **"Az emberi szervek vetítésétől a mesterséges ember vetítéséig" címmel.** Ernst Kapp 19. századi elképzeléséből indulva, miszerint az eszközök emberi szervek vetülései, azt állítják, hogy a technológiai fejlődés megfordította a folyamatot: az emberiség most már technológiailag rekonstruálhatja és kicserélheti magukat a szerveket. A tudatos kiborgizáció és a robotika végül egy új evolúciós szakaszt jelenthet.

Ez sokkal erősebb tézis, mint az a közismert állítás, miszerint "a technológia befolyásolja a társadalmat." A sorrend a következő:

szerv → szerv → szerv mesterséges helyettesítése → a szervezet → potenciálisan mesterséges emberi lény technológiai rekonstrukciója alapján modellezett.

Az MI ebben a sorozatban a legfontosabb emberi funkció, az **intelligencia technológiai rekonstrukciójaként lép** be.

2. "Cyborgoetika": Greuric legjellegzetesebb elméleti hozzájárulása

A normatív válasz, amit ő **kiborgoetikának** (*kiborgoetika*) nevez.

Célja, hogy szabályozza a legális technológiai beavatkozás és az átalakulások közötti határt, amelyek megszüntethetik azt a témát, amelyet állítólag javítanak.

2014-es szövege rendkívül messzemenő kérdéseket tesz fel: az emberiségnek engedélyeznie kellene-e testek, gének és agy módosítását; hogy a gondolatszabadság fennmarad-e, ha egy másik szereplő képes irányítani az agyat; hogy a fejlesztés elpusztíthatja-e az örökölt etikai értékeket; hogy mesterséges gondolkodású humanoidokat kell-e létrehozni; hogy az embereket "szuperemberekké" kell-e átalakítani; és hogy a korlátlan élettartamhosszabbítás egyáltalán felismerhetően emberi létezés maradna.

Greguric ezért azt szeretné, ha az etikai elvek végül jogi normákká válnának. Javasolt keretrendszere megkülönbözteti a terápiás helyreállítást a már normális képességek fejlesztésétől, megköveteli a biológiai korlátok és társadalmi következmények értékelését, multidiszciplináris és közvélemény-döntéshozatalt követel, és ragaszkodik hozzá, hogy a szabályozásnak meg kell védenie az egyéneket a katonai, vállalati és más hatalmas érdekektől. Meglepően előrelátó kérdése, hogy mi történik, ha a releváns "döntésközpont" végül maga a technológiává válik.

Ez fontos későbbi MI-elméletének megértéséhez. **Az MI összehangolása, az algoritmikus irányítás és az autonóm gép felelőssége az ő rendszerében egy általánosabb probléma speciális esetei: hogy a technológiailag megalkotott rendszerek alárendeltnek maradnak-e az emberi céloknak.**

Így Greguric elmélete implicit háromszintű normatív hierarchiával rendelkezik:

1. Technológia mint terápia: nagyrészt jogos, ha visszaállítja az elveszett funkciókat.

2. Technológia mint fejlesztés: etikailag vitatható, mert megváltoztatja az emberi alapvonalat.

3. Technológia mint autonóm utód: potenciálisan civilizációt megváltoztathat, mert az emberek talán nem lehetnek a technológiai fejlődés végső alanyai.

---

3. A 2018-as monográfia: a technológia egy egész társadalmi rend lesz

A döntő rendszerépítési munka Greguric 2018-as horvát monográfiája:

Kibernetička bića u doba znanstvenog humanizma: Prolegomena za kiborgoetiku
.

A könyv jelentősége pontosan a terjedelmében rejlik. Egy kortárs kritika szisztematikus áttekintésként és strukturált bevezetésként írta le a kiborg etikába, amely tizenhárom fejezetből áll.

Későbbi angol nyelvű kidolgozása, ***a Philosophical Issues of Human Cyborgization and the Necessity of Prolegomena on Cyborg Ethics*** különösen átláthatóvá teszi az építészetet. Fejezetei a "Tudományos humanizmus mint naturalizmus" és a technológia lényege a "Kiborg korszakán", "Hacking Human Beings", avatárok és kiborg antropológián át terjednek, majd **a "Valódi társadalmi kapcsolatok kiborgizálása", robotok, futúrológiai világrendek, kiborgoetika és végül egy javasolt általános határozat kiborgizációról** szólnak.

Ez az egyik legerősebb ok, amiért Greguric a közép- és kelet-európai szerzők közé sorolhatnám, akik átfogó **elméleti értelmezést próbáltak a fejlett technológia és társadalom** értelmezésére.

Nem csupán egyik etikai problémáról a másikra mozdul. Megpróbál átalakulásokat kialakítani a következőkben:

az emberi természet, a test, az intelligencia, az identitás, a társadalmi kapcsolatok, a kultúra, a kommunikáció, a törvény, az erkölcs és végül a civilizáció

egy alapvető történelmi átalakulásból.

És ez az alapvető átalakulás a **kibernetizáció**.

---

4. "Tudományos humanizmus mint naturalizmus": legmélyebb filozófiai tézise

Greguric munkájának legnehezebb, de intellektuálisan fontos eleme a **"tudományos humanizmus mint naturalizmus" fogalmának alkalmazása.**

Ezt nem szabad hétköznapi szekuláris humanizmusként értelmezni. Greguric érvelése egy filozófiai genealógiára támaszkodik, amely Hegel, Marx, Heidegger, kibernetika és horvát filozófiai értelmezések között szerepel a technológiai modernitásról. A kulcsfontosságú fejlemény az, hogy a modern tudomány és technológia már nem csupán *érti* a természetet; egyre inkább **létrehozza, szervezi át és tervezi át – beleértve magát az emberi** természetet is.

Ez a tézis rendkívül egyértelművé válik a 2020-as Springer-fejezetében **, "Transhumanizmus és mesterséges intelligencia: filozófiai aspektusok" című fejezetében.**

Ott Greguric azt írja, hogy a kibernetika egyfajta **új ontológiai tudomány lesz**. Az alkalmazott tudomány és technológia nemcsak új gépeket hoz létre, hanem egy új antropológiát, kozmológiát, és talán egy új teológiát is, amely egy technikailag átalakult világot idéz be, amelyben transzhumán kiborgok és poszthumán szuperintelligens robotlények is jelen vannak.

Valószínűleg ez az egyetlen mondat, amelyből az egész MI-elmélete a legjobban rekonstruálható.

Klasszikus technológiai társadalom

Az ember továbbra is az alany; a technológia egy eszköz; a természet a tárgy.

Kibernetikus társadalom

Az ember, a gép és a környezet egyre inkább összekapcsolt információfeldolgozó rendszerekké válnak.

Transhumanista társadalom

A technológia szándékosan változtatja meg az ember biológiai és kognitív alkotmányát.

Poszthumán társadalom

A biológiai emberi aljak történelmi kiváltsága eltűnik, mert intelligens mesterséges lények – vagyis hibrid ember-gép lények – maguk is cselekvés alanyává válhatnak.

Így a Greguric számára a végső MI forradalom **ontológiai, nem pusztán megrovás**.

---

5. MI és kibertér: az identitás átalakulása

Egy második jelentős összetevő a 2018-as Nenad Vertovšekkel közösen írt cikkben szerepel, **"Philosophy of Future Cyberspaces and the Transhumanistic Reality" címmel.**

Itt az elemzés a testből az információs környezet felé terjed.

Kiindulópontja, hogy a digitális valóság, a kiborgok, robotok és más mesterséges ügynökök alapvetően átalakítják az észlelést, a közvetítést és az emberek térhez való viszonyát. A transzhumanizmus ezt tovább viszi azzal, hogy szándékos irányítást javasol az emberi evolúció felett, és olyan lények létrehozását, amelyek intelligenciaban felülmúlják a jelenlegi embereket.

Ez létrehozza azt, amit Greguric **identitásproblémájának nevezhetnénk az MI civilizáció** területén.

Az emberi identitás hagyományosan egy többé-kevésbé stabil kapcsolatot feltételez a következők között:

test, elme, emlékezet, hely, ügynökség és személyes múlt.

De mind a hat technológiailag változóvá válhat.

Egy személy mesterséges szerveket is birtokolhat; a kognitív implantátumok módosíthatják az emlékezetet vagy észlelést; az interakció egyre inkább avatárokon keresztül történhet meg; a testi implantátumok által generált információk vállalati hálózatokban lehetnek; algoritmikus rendszerek közvetíthetik azt, amit az egyén lát; és az intelligens gépek maguk is egyre inkább ügynökszerű funkciókat szerezhetnek.

Ennek következtében a kérdés már nem csupán az, hogy "Hasznos-e ez a technológia?" Ez így alakul:

Hol ér véget az egyén, és hol kezdődik a technológiai rendszer?

Greguric és Vertovšek ezért egy olyan kérdést vet fel, amely a cikkük óta még fontosabbá vált: **ki birtokolja és irányítja a technológiailag kibővített testekben generált információt?**

Ez egyesíti a bioetikát, az adatvédelmet és az MI-irányítást egyetlen problémába.

---

6. Információs hatalom: a megfigyeléstől az észlelés irányításáig

Greguric elemzése fontos politikai dimenziót tartalmaz, amelyet nem szabad figyelmen kívül hagyni.

A 2018-as kibertér tanulmány olyan forgatókönyveket tárgyal, amelyekben a digitálisan közvetített észlelés erősen személyre szabott és szűrt lehetne. Bár néhány konkrét előrejelzés – például Ray Kurzweil 2030 körüli kiterjedt agy-számítógép hibridizáció előrejelzése – olyan futuristák előrejelzései, amelyeket Greguric tárgyal, nem pedig Greguric saját előrejelzései, ő maga is politikai következtetést von: **az információszűrés feletti irányítás átvihetne vállalatokra, kormányokra vagy más szervezett érdekekre**, így valamilyen "infoetika" szükséges lenne.

2025-ös kommunikációs elemzése jelentős összefüggést mutat ezzel a régebbi érvel, de ez a mai ajánlási algoritmusok és generatív mesterséges intelligencia világába is átültethető.

Greguric azt állítja, hogy a digitális platformok már nem csupán továbbítják az információt. Az algoritmikus mediáció befolyásolja, hogy az emberek szembesülnek, és ezért az ő fogalmazása szerint egyre **inkább azt is befolyásolja, hogyan gondolkodnak**. Az elköteleződést maximalizáló algoritmusokat összekapcsolja a polarizációval, manipulációval, információbuborékokkal és egy közös információs környezet eróziójával.

Ez egy másodfajta emberi felerősítést – vagy esetleg emberi csökkenést – vezet be az elméletébe:

a fizikai kiborgizáció integrálja a gépeket a testekbe;

Az információs kiborgizáció algoritmikus rendszereket integrál az észlelésbe és a kognícióba.

A második beültetés nélkül is előfordulhat. Az a személy, akinek a társadalmi valóság észlelését folyamatosan ajánlási algoritmusok választják, már részt vesz egy kibernetikus visszacsatolási rendszerben.

Ez fontos kiterjesztése Greguric eredeti elméletének.

---

7. Mesterséges intelligencia mint autonóm társadalmi szereplő

A legradikálisabb mesterséges intelligencia szöveg valószínűleg a 2020-as cikk Vertovšekkel,

"Filozófia és tudatosság a jövőben: Kiborgok és mesterséges intelligencia várnak a halhatatlanságra."

Itt az elemzésük határozottan túllép a kiborgokon és **maga az MI önálló ügynökségi központtá** válik.

Azt állítják, hogy a mesterséges intelligencia, a robotika és a kibernetikus rendszerek a civilizáció központja felé haladnak, miközben a virtuális és kiterjesztett valóságok alapvetően megváltoztathatják a kommunikációt és a társadalmi létezést. A tanulmány még azt is állítja, hogy az MI elkezdte kialakítani saját tudatát, és végül kidolgozhat egy fejlesztési stratégiát "az emberiségen túl" vagy "nélkül".

Ez az állítás kritikus minősítést igényel.

Nem szabad empirikusan megállapított megállapításként értelmezni, hogy a jelenlegi MI rendszerek tudatosak. A cikk nem nyújt empirikus tudattesztet, amely ezt bizonyítaná. Jobban értelmezhető filozófiai-futurológiai extrapolációként: a növekvő összetettség és autonómia végül megkérdőjelezheti azt a feltételezést, hogy az intelligencia és a tudatosság biológiailag emberiek maradjanak.

Más szóval, három állítást kell különválasztani:

Az MI autonómia – a rendszerek folyamatos emberi beavatkozás nélkül is képesek egyre jelentősebb döntéseket hozni – technológiai tendenciá.

A mesterséges általános vagy szuperintelligencia a jövőben is lehetséges.

A mesterséges tudatosság sokkal erősebb filozófiai állítás, és nem pusztán a számítási képességek növelésével állapítható meg.

Greguric és Vertovšek néha gyorsan átmozognak ezek között a szintek között. Ez az egyik fő gyengesége a jövőbeli érvelésüknek.

---

8. A munka átalakulása

Ugyanebben a 2020-as tanulmányban felemelkedően radikális tézis is található a foglalkoztatásról.

A szerzők a robotokat és az intelligens rendszereket az emberi "mechanikus leszármazottaknak" képzelik el, akik fokozatosan képesek mind **fizikai, mind szellemi munkát** vállalni. Ezt a biológiai és technológiai intelligencia és ügynökségi formái közötti nagyobb átmenetbe helyezik.

Itt azonban fontos meghatározni, hogy a gregurikus nem kifejtő .

Nem nyújt részletes munka-gazdasági modellt, amely összehasonlítható lenne a kortárs kutatásokkal, amelyek a foglalkozási kitettséget, a bérhatásokat, a termelékenységet, a munkanélküliséget vagy az újraelosztást becslékék. És nem épít fel részletes elméletet az AI alatti osztályátalakításról.

Hozzájárulása alapvetőbb, de kevésbé empirikusan meghatározott jellegű:

Amint az intellektuális tevékenység technológiailag reprodukálhatóvá válik, a munka már nem kizárólag emberi területet alkot.

Ez filozófiailag radikális. Szociológiailag azonban további elméletre van szükség, hogy elmagyarázza, ki birtokolja az MI rendszereket, ki részesül termelékenységnövekedést, hogyan szerveződnek át a szakmák, hogyan reagálnak a jóléti államok, és hogy a helyettesítő vagy kiegészítő augmentáció dominál-e.

Így a **munkaerő kiszorítása az általános elméletében létezik, de nem a legfejlettebb alrendszere**.

---

9. Vállalati hatalom – és egy fontos korrekció az egyszerű "MI átvétel" narratívákhoz

Valójában érdekes feszültség van Greuric írásaiban.

Egyrészt azt fontolgatja, hogy az autonóm vagy szuperintelligens MI kiszabadulhat az emberi irányítás alól.

Másrészt a 2020-as tanulmány azt állítja, hogy közvetlenebb veszélyt jelent, ha **az emberek saját érdekeik érdekében használják** az MI-t. Tárgyalja a digitális platformokat, érdekcsoportokat és technológiailag erős szervezeteket, és lényegében arra a következtetésre jut, hogy a veszély kevésbé a gépi intelligenciában, inkább az emberi visszaélésben – vagyis az "emberi nem-intelligenciában" rejlik.

Ez két különböző technológiai dominancia modellt eredményez:

A modell: az emberek az MI-n keresztül uralják az embereket.

Vállalatok, kormányok, katonai szervezetek vagy technológiai elit irányítják az algoritmusokat, információkat, fejlesztési technológiákat és adatokat.

B modell: Az MI elmenekül az emberi uralom elől.

Egyre autonóm rendszerek olyan képességeket vagy célokat szerzetek, amelyeket az emberek már nem tudnak hatékonyan felügyelni.

Greuric keretrendszere mindkettőt tartalmazza.

Ez elméletileg értékes, mert megakadályozza, hogy az AI kockázata gépek által sci-fi lázadássá változzon. **Mielőtt a technológia autonóm hatalmi központtá válhatna, a fejlett technológia birtoklása maga is drámaian újraoszthatja a hatalmat az emberek között.**

---

10. Orvostudomány, fejlesztés és mesterséges intelligencia: a gyógyítástól az áttervezésig

Az orvosi dimenzió az egyik legerősebb része az elméletének.

A Greguric ismételten megkülönbözteti azokat a technológiákat, amelyek **helyreállítják a sérült emberi kapacitást**, és olyan technológiák között, amelyek **az egészséges egyének képességét a normál emberi hatótávolságon túlra** növelik.

Egy betegség utáni funkciót helyreállító idegi implantátum, valamint egy egészséges ember kognitív képességének növelésére használt idegimplantátum nem kaphat automatikusan ugyanazt az etikai státuszt.

Ez a látszólag egyszerű megkülönböztetés óriási következményekkel jár.

Amint az MI-fejlesztéssel fejlesztett agy-számítógép interfészek eléggé kifinomultak, az orvosi technológia három szakaszon haladhat át:

kezelés → bővítés → áttervezés.

Az első megőrzi az embert.

A második megváltoztatja az emberi képességeit.

A harmadik potenciálisan megváltoztatja az emberiség definícióját.

Greguric kiborg etika és robotetika kurzusa kifejezetten integrálja az NBIC konvergenciát – nanotechnológiát, biotechnológiát, információs technológiát és kognitív tudományt – az MI-vel, a transzhumanizmussal, agymódosítással, robotokkal, szuperintelligenciával, magánélettel, felelősséggel és jogi szabályozással.

Ennek a keretrendszernek a szélessége az egyik oka annak, hogy munkáját sokkal közelebb tartom a technológiai társadalom szisztematikus filozófiájához, mint a hagyományos alkalmazott MI-etikához.

---

11. Maga a társadalom "kibernetikus organizmusná" válik

Greuric egyik legelméletileg ambiciózusabb elképzelése a **"tényleges társadalmi kapcsolatok kiborgizálása"** és a társadalom kiborgizálódása, mint kibernetikus organizmus. Ez kifejezetten megjelenik az építészeti tanításában és monográfiáiban.

Az ötlet jelentősége a következőképpen rekonstruálható.

A kibornetika információn, visszacsatoláson, irányításon és alkalmazkodáson alapul.

A modern digitalizált társadalmak egyre inkább pontosan ilyen mechanizmusokon keresztül működnek:

Az emberek adatokat generálnak, → algoritmusok dolgozzák fel azokat, → a rendszerek módosítják a környezetet és ajánlásokat, → az emberek reagálnak, → új viselkedési adatok keletkeznek, → a rendszerek újra tanulnak és beavatkoznak.

Az eredmény egy hatalmas visszacsatolási struktúra.

Így Greguric hatékonyan áthelyezi a kiborg metaforáját az egyéni testről a társadalomra:

Egyéni kiborg: biológiai organizmus + technológiai visszacsatolási rendszerek;

Kibernetikus társadalom: emberi népesség + hálózatok + algoritmusok + gépek + adatbázisok + automatizált visszacsatolás.

Ez az egyik olyan pont, ahol az elmélete a legközelebb kerül az igazán átfogó társadalomelmélethez.

---

12. Kultúra, művészet, sport, közlekedés és mindennapi élet

A Greuric rendszerének másik erőssége, hogy a kiborgizáció nem korlátozódik laboratóriumokra.

Elméleti programja kifejezetten alkalmazza a fogalmat sportra **, művészetre, zenére, közlekedésre, virtuális valóságra és hétköznapi társas kapcsolatokra**. Tanítási

keretrendszere a technológiailag fejlett sportolókról, a technológiailag módosított testekről a performanszművészetben, a szimulált valóságról, a kibernetikus szállításról, a szociális robotokról és végső soron az MI-ről szól.

Az alaptézis az, hogy a határprobléma mindenhol visszatér.

Mikor kompenzálja a mesterséges protézis a fogyatékosságot, és mikor teremt versenyelőnyt?

Mikor válik a technológia művészeti médiummá, és mikor válik maga a technológiai módosítás a műalkotásnak?

Mikor segíti a navigációs szoftver a vezetőt, és mikor válik az autonóm rendszer a vezetővé?

Mikor támogatja a szoftver a gondolkodást, és mikor hajtja végre a kognitív műveletet helyettünk?

Ugyanez a fogalmi struktúra tehát működik a protézisektől a generatív MI-g.

---

13. Oktatás: az elmélet újabb bővítése

Az oktatás viszonylag új explicit kiterjesztés.

A **Cambridge-i Generatív MI az Oktatásban 2024-es** konferencián Greguric *bemutatta "Transhumanizmus és az emberi mesterséges intelligencia jövője: filozófiai következmények az oktatásra és identitásra."* Az előadás összekapcsolta a technológiai fejlesztést és az ember–mesterséges intelligencia interakcióját a jövőbeli oktatási gyakorlatokkal és az identitás változásaival, és kifejezetten megvizsgálta a kiborgizáció optimista és romboló értelmezéseit, valamint a "harmadik út" lehetőségét.

Ez a "harmadik út" fontos az ő általános helyzetének értelmezéséhez.

Nem **szabad egyszerűen biokonzervatív technológiaellenfélként besorolni.**

De ugyanolyan messze van a kritikamentes transzhumanizmustól.

Pozíciója közelebb áll hozzá:

Elfogadjuk a terápiás és emberileg hasznos technológiákat + ismerjük fel a hatalmas felszabadító lehetőségeket + állítsunk fel szigorú etikai határokat, ha a fejlesztés veszélyezteti az autonómiát, identitást vagy emberi cselekvést.

Ennek következtében az ember–mesterséges intelligencia kapcsolatot irányítani kell, nem egyszerűen elutasítani vagy korlátok nélkül fejlődni.

---

14. MI, újságírás és az igazság válsága

Greguric 2025-ös kortárs média témájú tárgyalása sokkal konkrétabbá teszi korábbi filozófiáját.

Leírja a hagyományos információs közvetítéstől a töredezett platformok, influenszerek, személyre szabott tartalom és algoritmikus kiválasztás világa felé vezető mozdulatot. Az MI által generált újságírás egy régebbi problémát erősít: mit jelent az "igazság", amikor szövegek generálhatók anélkül, hogy emberi tanú vagy szerző átélné az eseményt?

Javasolt válaszai feltáróak:

erősebb kritikus gondolkodás,

médiaműveltség,

nagyobb társadalmi ellenálló képesség a manipulációval szemben,

valamint a digitális platformok szabályozása.

Itt az MI már nem elsősorban egy képzeletbeli jövőbeli szuperintelligencia. Már **a társadalom episztemikus infrastruktúrájának része**.

Így jelentős fejlemény történik Greguric gondolkodásában:

korai Greguric: a technológia átalakíthatja az emberi testet;

középső gregur: kibernetikus technológia megváltoztathatja az emberi természetet;

2020 Greguric: Az MI autonóm intelligenciaformává válhat;

A legutóbbi Greguric: algoritmusok és generatív MI már most is átalakítják a társadalom termelését és valóságérzékelését.

---

15. Az AI közelgő radikális szakaszainak elmélete

Greuric munkásságának jövőbeli összetevője most már pontosabban is megfogalmazható.

Nem egyetlen technológiai forradalomot, hanem **több összesített küszöböt képzel** el.

Első küszöb: mindenütt jelenlévő algoritmikus közvetítés

Ez már most is megtörténik. Az MI közvetíti az információt, a kommunikációt, a döntéshozatalt, a médiagyártást, valamint egyre inkább az oktatást és tudásmunkát. Greguric

legutóbbi médiaelemzése a személyre szabást, a fragmentációt, a félretájékoztatást és az algoritmikus befolyásolást hangsúlyozza az észlelés felett.

Második küszöb: szociális robotok és autonóm mesterséges ügynökök

A robotok már nem csupán ipari gépek, hanem egyre inkább belépnek hétköznapi emberi környezetbe. Greuric jelenlegi tanterve kifejezetten a robotokkal való együttélésről, autonómiáról, bizalomról, magánéletről és felelősségről kérdez, amikor az autonóm rendszerek ártanak az embereknek.

Harmadik küszöb: a biológiai és mesterséges intelligencia integrációja

Az MI kombinálódik neurotechnológiával, implantátumokkal, protézisekkel és kognitív fejlesztésekkel. Ezen a ponton már nem csupán az emberek *MI-t használják* .

Ez az **emberi intelligenciává válik, amelyet részben mesterséges rendszerek alkotnak**.

Ez a modellje valóban transzhumán fázisa.

Negyedik küszöb: mesterséges általános vagy szuperintelligencia

Greguric komolyan veszi a szuperintelligencia lehetőségét. Egy 2025-ös tudományos kolokviumon Bostrom szuperintelligencia forgatókönyveiről beszélt, beleértve az irányítás elvesztését és a rendszer önreplikációját, miközben hangsúlyozta az erős etikai, jogi és társadalmi keretek szükségességét.

Ötödik küszöb: poszthumán civilizáció

Modellje legradikálisabb végén a biológiai emberiség abbahagyhatja az intelligencia és a történelmi cselekvési rendszer monopolizálását. Kibernetikus lények, autonóm robotok, avatárok és mesterséges intelligencia egy újfajta társadalmi világot alkothatnának.

Ez az átmenet egy **MI-t használó emberi civilizációból egy vegyes vagy poszthumán civilizációba**.

---

16. Milyen dátumokat jósolnak meg valójában a Greguric?

Fontos bibliográfiai különbség van itt, mert Greuric vitákban megjelenő drámai dátumok **nem saját jóslatok**.

Például 2018-as cikke **Ray Kurzweil** elvárásait tárgyalja a radikális agy-gép hibridizáció 2030 körüli elvárásáról.

Munkája más futuristák szingularitásáról, feltöltéséről és technológiai halhatatlansági helyzeteiről is tárgyal. Valóban, későbbi monográfiájának egyik fejezete kifejezetten

"Futurológus előrejelzések a kiborgok és robotok globális világrendjéről" címet viseli. A kiadó úgy írja le a fejezetet, hogy olyan futuristák jóslatait mutatja be, akik az emberi fejlődés és a technológiai átalakulás iránt foglalkoznak.

Ezért helytelen lenne az összes ilyen dátumot közvetlenül a gregurhoz kötni.

Saját előrejelzései általában **irányszerűbbek és** feltételesebbek.

A legerősebb példa a 2014-es tanulmányában található, ahol azt írja, hogy **a huszo1. század végén az emberiség bionikus testekkel rendelkezhet, mesterséges agyat és érzékszerveket**, és feltételesen érvel, hogy a folyamatos fejlesztés részben mesterséges emberekről – kiborgokról – teljesen mesterséges lények felé léphet.

Ez valóban az ő radikális extrapolációja, de ez egy **forgatókönyv**, nem egy empirikusan kalibrált előrejelzés.

---

17. A "szingularitás" Greguric elméletében

A Greguric szingularitással kapcsolatos irodalmat használ, de szélesebb filozófiai jelentést ad neki, mint egy egyszerű "számítógép okosabb lesz egy embernél" eseményen.

A rendszerében a valódi szingularitás több határ egyszerre összeomlásából állna:

ember/mesterséges intelligencia,

természetes/mesterséges test,

szervezet/gép,

valódi/virtuális környezet,

emberi/nem emberi cselekvési lehetőség,

eszköz/felhasználó,

és talán végül biológiai/posztbiológiai életet.

Ez teszi a fogalmát sokkal érdekesebbé szociológiailag, mint egy egyszerű számítási teljesítménygrafikon.

Egy civilizáció akár egy gregur-stílusú poszthumán átalakuláson **is megeshetett volna egyetlen pontosan datálható szingularitási esemény** nélkül, mert az MI, robotika, biotechnológia, hálózatok és emberi test fokozatos integrációja fokozatosan feloldhatná a régi kategóriákat.

---

18. Az MI veszélye: a kihalás csak egy lehetőség

Elméletének másik hasznos jellemzője a technológiai kockázat sokkal szélesebb körű felfogása.

A szélsőséges forgatókönyv nyilvánvaló: az autonóm szuperintelligencia egyre kontrollálhatatlanná válik.

De Greuric mélyebb aggodalma az **átalakulás, nem pedig pusztán a fizikai kihalás**.

Az emberiség biológiailag túlélhette volna, miközben elveszítette a történelmi emberiséghez kapcsolódó fontos jellemzőket:

autonóm ítélet,

testi integritás,

gondolati magánélet,

stabil személyes identitás,

a személyes adatok feletti ellenőrzés,

közvetítetlen hozzáférés a társadalmi valósághoz,

vagy maga a technológiai fejlődés feletti irányítás.

Ez egyfajta **antropológiai kiszorításnak nevezhető** formája.

A gépnek nem feltétlenül kell elpusztítania az emberiséget. Az emberiség maga is fokozatosan újratervezhetné intézményeit, testét és kognitív folyamatait, míg az eredeti különbség az ember és a gép között nem bír jelentőséget.

Ez finomabb állítás, mint a szokásos MI-apokalipszis forgatókönyv.

---

19. Politikai és jogi következmények

Greguric normatív válasza szokatlanul intézményes egy technológiafilozófus esetében.

Már 2014-ben azt állítja, hogy az etikai következtetéseket a fejlesztéssel kapcsolatban végső soron jogi **szabályokban kell** megtestesíteni, szoros nyilvános ellenőrzés alá vonva, különös tekintettel a katonai és vállalati hatalomra.

Jelenlegi kurzusa kiterjeszti a problémát a robotetikára is, és azt kérdezi, ki vagy mi legyen jogilag vagy erkölcsi felelősséggel, ha autonóm rendszerek kárt okoznak.

Egy érett görög típusú szabályozási rendszernek tehát legalább négy különböző kérdésre kell választ adnia:

ki fejleszthet technológiát; ki alkalmazhatja azt; milyen emberi módosítások engedélyezettek; és ki vállalja a felelősséget, amikor technológiailag autonóm rendszerek jelentős döntéseket hoznak.

Ez egy átfogóbb szabályozási fogalma, mint a szokásos "MI biztonság", mert a szabályozott objektum az **egész ember–gép kapcsolat**.

---

20. Pozíciója a transzhumanizmus és a biokonzervativizmus között

Ez talán a legfontosabb értelmezési pont.

Greguric alaposan tanulmányozza a transzhumanistákat, mint Kurzweil és más futuristák, de **nem csupán támogatja a programjukat**.

Nem támogatja a technológiai fejlesztés megállítását sem.

Alapvető kritikája a radikális transzhumanizmussal szemben, hogy az erősítés belső paradoxont tartalmaz:

Ha valaki korlátlanul fejleszti az embert, egy ponton az eredmény már nem egy fejlődő ember, hanem egy másfajta lény.

Ezért ragaszkodik Greguric a határkérdéshez.

A kiterjesztés nem feltétlenül marad meg, mivel az átalakulás nagysága a végtelen felé hajlik.*HH*

Ez a kiborgoetika **logikai magja**.

A gyakorlati nehézség természetesen az, hogy meghatározzák, hol helyezkedik el ez a határ.

És Greguric nem ad egyszerű numerikus választ. Ehelyett egy megfontoló, etikus, orvosi, technológiai, filozófiai, jogi és nyilvános folyamatot javasol.

---

21. A 2021/2023-as szintézis: a "nem emberi" filozófiája

További szakasz következik Greguric és Vertovšek ***Filozofija budućeg: Ogledi o neljudskom*** (2021) című műve, amely később angolul megjelent ***Philosophies of the Future and the Non-Human: From Cyberspace to Human Enhancement*** (2023) címmel.

A kiadó kifejezetten úgy írja le célját, hogy megvizsgálja, mi számít "embernek lenni" a gépekkel és az MI-vel való együttélés feltételei között, és azt mondja, hogy a szerzők egy olyan tudományterületet céloznak, amely egy technológiai jövőt elemz, amelyet fejlett kiborgok, avatárok és autonóm robotok laknak.

Az angol nyelvű könyv a 2021-es horvát kötet fordítása, nyolc esszét és egy interjút Stelarc-szal tartalmaz.

Ez a könyv fontos fogalmi bővülést jelent.

A probléma már nem csupán:

Mit fog tenni a technológia az emberekkel?

Ez így alakul:

Milyen lények fognak olyan jövőben élni, ahol az "ember" már nem az egyetlen releváns kategória?

Ez az átmenet – a technológiai etikáról a nem emberi filozófiára – vitathatatlanul Greguric elméletének legradikálisabb fejleménye.

---

22. Egy fő elméleti erősség: összekapcsolja a külső mesterséges intelligencia és a belső emberi fejlesztés

Ez megkülönbözteti a gregurikus művet a kortárs MI irodalom nagy részétől.

Az MI-elmélet gyakran vizsgálja **a külső gépi intelligenciát**:

emberi ↔ MI.

A transzhumanizmus az **emberi intelligencia fejlesztését** vizsgálja:

emberi → felerősített ember.

Greguric csatlakozik a kettőhöz:

ember + MI + biotechnológia + neurotechnológia + robotika → kibernetikui/poszthumán ügynökségi formák.

Ez azért fontos, mert a két technológiai pálya összeegyezhet.

Egy olyan jövő, ahol az MI sokkal erősebb, de az emberektől kívül marad, egészen más attól, ahol az emberek MI-hez kapcsolódó idegi interfészekkel, mesterséges memóriával, kibővített észleléssel és folyamatos gépi közvetítéssel rendelkeznek.

Greuric elmélete különösen erősen előrevetíti a második lehetőséget.

---

23. Egy másik erősség: a technológiai evolúció egyidejűleg társadalmi evolúció is

2021-ben Džinić-csel közösen írt cikke ezt különösen világossá teszi. Az emberi szervek technológiai rekonstrukcióját nemcsak orvosi fejlődésként értelmezik, hanem akár egy **új evolúciós szakasz potenciálisan** hozzájárulóként.

Ez szokatlan kölcsönhatást eredményez három evolúciós folyamat között:

a biológiai evolúció létrehozta a Homo sapiens-t;

a kulturális evolúció nyelvet, intézményeket és felhalmozódott tudást hozott létre;

a technológiai önevolúció lehetővé tehetné a Homo sapiens számára, hogy szándékosan újratervezze önmagát, és nem biológiai utódokat hozzon létre.

Az MI döntő a harmadik szakaszban, mert először a technológiailag újraépítendő kapacitás az a képesség, amely maga a technológiai rekonstrukció felelőse: az **intelligencia**.

Ez lehetőséget teremt egy rekurzív folyamat kialakulására.

Az emberek MI-t hoznak létre → az MI segít jobb technológia létrehozásában, → a jobb technológia javítja az embereket és a gépeket, → a fejlett intelligencia felgyorsítja a további technológiai fejlődést.

Ez a rekurzív logika Greguric radikális gyorsulás iránti aggodalmának egyik implicit alapja.

---

24. De a jövővizsgálatának fontos gyengeségei vannak

Elmélete intellektuálisan ambiciózus, de több korlátot is tisztán szem előtt kell tartani.

Először is, az **MI, robotika, kibernetika, biotechnológia, virtuális valóság és emberi fejlesztés néha egyetlen konvergens történelmi mozgalom összetevőiként kezelik**. Ez nagy

szélességet ad elméletének, de elhomályosíthatja azt a tényt, hogy ezeknek a technológiáknak nagyon eltérő fejlődési pályájuk és társadalmi következményei vannak.

Másodszor, néhány legradikálisabb technológiai állítása nagymértékben **Kurzweilre, Bostromra, Kakura és más futuristákra** támaszkodik. A kifejezetten futurológusoknak szentelt fejezet ezt megerősíti. Ezek a forgatókönyvek filozófiai felépítésében használt anyagok; nem mindegyiket szabad önállóan megalapozott gregurikus előrejelzésként bemutatni.

Harmadszor, a 2020-as tanulmányban szereplő állítás, miszerint az MI már elkezdte fejleszteni saját tudatát, **sokkal erősebb, mint a cikkben bemutatott érv**. Ezt inkább spekulatív filozófiai tézisként kell kezelni, nem pedig bizonyított eredményként.

Negyedszer, modellje kissé **teleológiai felfogásra hajlik a technológiai fejlődésről**: a folyamatos konvergencia úgy tűnik, hogy kiborgokra, poszthumánokra és mesterséges intelligenciára utal. A tényleges technológiai történelmet megszakíthatják gazdasági költségek, szabályozás, kulturális ellenállás, mérnöki korlátok, háborúk, intézményi választások és versengő technológiák.

Ötödrész, az elmélet sokkal erősebb az **antropológiában és ontológiában, mint a politikai gazdaságtanban**. Kevésbé szisztematikus kezelést ad az MI tulajdonlásáról, tőkekoncentrációról, adózásról, jóléti rendszerekről, szakmai átszervezésről, egyenlőtlenségről, állami kapacitásról és geopolitikai versenyről, mint amennyit egy valóban átfogó szociológiai elmélet megkövetelne az MI-ről.

Ezek a korlátok nem rombolják el az elméletet. Pontosan megmondják, milyen elméletet épített fel Greguric.

---

25. A spekulatív filozófiától az empirikus MI szociológiáig: az új AI-COM projekt

Különösen jelentős újonnan történt fejlesztés a Zágrábi Egyetem **"Mesterséges Intelligencia: Horvát Társadalom Új Közvetítője" (AI-COM)** projektje, amelynek Greguric kutatója.

A négyéves projekt jelentősen szélesíti az empirikus napirentereket. Lefeszegeti az MI filozófiai alapjait; az MI alkalmazását a társadalomban és annak üzleti hatását; MI a tudományban és a PR-ben; kreatív MI-t; valamint az etikai normákat. Célja, hogy ötvözze a történelmi és elméleti elemzést minőségi és mennyiségi kutatással, tanulmányozza, hogyan értik és használják az MI-t, valamint azonosítva azokat a csoportokat, akik különösen hátrányos következményekkel vannak kitéve.

A projekt első munkatalálkozója 2026. április **1-jén zajlott**, ahol Greguric egy filozófia, kommunikációtudományi, szociológia és pszichológia interdiszciplináris csapata tagja volt.

Ez azért fontos, mert ez egy **új szakaszt jelenthet az intellektuális fejlődésében**:

2010-es évek: filozófiai-bioetikai fejlesztéselmélet;

2010-es évek vége: kibernetikus társadalom és poszthumán identitás;

2020–23: MI, szuperintelligencia és nem emberi jövők;

2024–25: generatív MI, oktatás, információs és média;

2026-tól kezdve: az AI tényleges társadalmi behatolásának szisztematikus empirikus vizsgálata.

Mivel az AI-COM csak nemrég kezdődött, a jövőbeli eredményeit még nem szabad Greguricnak tulajdonítani. De a tervezése pontosan az átfogó társadalmi elemzés felé vezető lépést jelez, amelyet a kérdésed hangsúlyoz.

26. Mi valóban eredeti Greguric elméletében?

Ha szigorú eredetiségi kritériumot alkalmazunk, **négy különösen jelentős hozzájárulást** választanék meg.

Az első a kiborgoetika mint rendszerszintű normatív terület.

Greguric nem csupán katalogizálja az etikai problémákat. Megpróbál egy koherens területet létrehozni, amely szabályozza a technológiai beavatkozás emberi korlátait, és ezeket az etikai elveket összekapcsolja a jövőbeli jogi szabályozással.

A második az átalakulás az "eszközetikától" az "ontológiává".

A technológia nem csupán a társadalom által használt dolog. A kibernetikus racionalitás egyre inkább meghatározza, **hogyan épülnek létre maguk az élet, az intelligencia, a testek és maguk a társas kapcsolatok**. Springer-megfogalmazása, amely a kibernetikáról új ontológiai tudományként alkotja, különösen jól összefoglalja ezt a lépést.

A harmadik az emberi fejlesztés és az MI egysége.

Ahelyett, hogy az MI-t és a transzhumanizmust különálló jelenségként tárgyalná, úgy értelmezi őket, mint az emberiség kibornetikus átalakulásának összefonódó útvonalainak.

A negyedik az antropológiai kitelepítés fogalma.

Legmélyebb félelme nem csupán a munkanélküliség vagy akár az ellenséges robotok. Az az, hogy a korlátlan fejlesztés és a mesterséges intelligencia végül megszüntetheti a **természetes ember kategóriáját, mint a történelem megkérdőjelezhetetlen** témáját.

Ez radikálisabb, mint a legtöbb mainstream AI-etikai keretrendszer.

27. Általános értékelés

Értékelésem szerint **Ivana Greguric valóban egy átfogó elméleti felfogás szerzője az MI jövőbeli társadalomban betöltött helyéről**, feltéve, hogy a "átfogó" filozófiailag érthető, nem pedig minden társadalmi alrendszer teljes empirikus szociológiájaként.

Elméleti felépítése a következő képletbe sűríthető:

A kibornetika → az élet technológiai rekonstrukciója → emberi fejlesztés → kiborgizáció → az MI autonómia → szuperintelligencia → posztemberi/nem emberi együttélés → a kiborgoetika és robotetika szükségessége.

A központi ok-okozati folyamat a **technológia fokozatos internalizációja és autonómizálása**.

Kezdetben a technológia az embert veszi körül.

Ezután belép a testbe.

Ezután belép a kognícióba.

Ezután az intelligens technológiai rendszerek együttműködnek az emberekkel.

Így önállóan hozhatnak döntéseket.

Végül, Greguric legradikálisabb forgatókönyvében önálló történelmi fejlődési alanyokká válhatnak.

Legmélyebb tézise tehát nem az, hogy **az MI csupán megváltoztatja a társadalmat**.

Az a lényeg, hogy az MI, a robotikával, a neurotechnológiával és a biotechnológiával kombinálva, végül megváltoztathatja **a társadalmat alkotó entitásokat**.

Ezért a foglalkoztatás, kommunikáció, orvostudomány, magánélet vagy szabályozás kérdései – bármennyire is fontosak – maradnak a Greuric rendszerének mélyebb átalakulásának másodlagos megnyilvánulásai. Az elsődleges átalakulás egy olyan civilizációból egy olyan civilizációba, amelyben **a biológiai emberek gépeket használnak**, olyan civilizációvá, ahol **a biológiai emberek, technológiailag fejlett emberek, kibornetikus hibridek és autonóm mesterséges intelligenciák különböző ügynökségi formákként együtt létezhetnek**.

Ez magyarázza továbbá kiborgoetika iránti felhívásának jelentőségét is: **szerinte a döntő politikai és etikai döntéseket még azelőtt kell meghozni, hogy maga a technológiai fejlődés határozza meg az ember határait.** Jelenlegi munkája a generatív MI, oktatás, média és az AI-COM projekt területén azt sugallja, hogy most ezt a hosszú távú filozófiai keretrendszert érinti az MI által a 2020-as években előidézett sokkal közvetlenebb társadalmi átalakulásokkal.

****

Nenad Vertovšek elmélete

Kérjük, részletesen írja le és elemezze Nenad Vertovšek tanulmányait , amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint látnak.

Nenad Vertovšek különösen 2018-tól egy felismerhető kutatási programot dolgozott ki a mesterséges intelligenciáról, kiberkultúráról, transzhumanizmusról, a médiáról, a tudatosságról és az emberi társadalom jövőbeli átalakulásáról. Kissé félrevezető lenne azonban technikai értelemben MI-kutatóként nevezni. Nem épít gépi tanulási modelleket, és nem készít kvantitatív előrejelzéseket a számítási teljesítményről. Jellegzetes hozzájárulása **filozófiai, médiaelméleti, szociális-elméleti és egyre inkább politikai-gazdasági** jellegű: azt kérdezi, mi történik a társadalommal, amikor az MI nem csupán technikai eszköz, hanem annak a környezetnek a részévé válik, amelyen keresztül az emberek érzékelik a valóságot, kommunikálnak, hatalmat gyakorolnak, dolgoznak, háborúznak, és végső soron meghatározzák, mit jelent embernek lenni.

Van egy fontos attribúciós pont is. A 2018–2023-as alapító munkák nagy részét a horvát **filozófussal, Ivana Greguric-csal** közösen írták, akinek saját szakterületei közé tartozik a kiborgizáció, a transzhumanizmus, a bioetika és a "kiborg etika". Ezért a poszthumán és kiborg dimenziót nem szabad kizárólag Vertovšeknek tulajdonítani. Különösen világosan meghatározza Vertovšek független álláspontját az ő 2025-ös szóló cikke*, az Izazovi umjetne inteligencije – ne-ljudska budućnost?!* valamint 2025–2026-os írásai és interjúi, amelyekben személyesen is a korábbi keretrendszert alkalmazza az MI, algoritmusok, vállalati hatalom és a "technofeudalizmus" elemzésébe.

1. Vertovšek MI-elméletének fejlődése

Releváns munkássága viszonylag koherens sorozatként rekonstruálható:

Időszak	Fő művek	Elméleti lépés
2018	Vertovšek & Greguric, *A jövőbeli kiberterek filozófiája és a transzhumanista valóság*	Kibertér, transzhumanizmus, emberi identitás és az emberi/nem emberi határ feloldása
2019	Vertovšek & Greguric, *Az ember és a világ tudományos-technikai és média-kapcsolódó jövője*	Az MI központi problémává válik; egy közelgő "fordulópont" a természetes és a mesterséges intelligencia között
2020	Vertovšek és Greguric Knežević, *Filozófia és tudatosság a jövőben: kiborgok és mesterséges intelligencia halhatatlanságra várva*	Gépi tudatosság, autonóm MI fejlesztés, kiborgizáció és technológiai halhatatlanság
2021/2023	*Filozofija budućeg: Ogledi o neljudskom* / *A jövő és az emberi nem filozófiák*	Könyvhosszú szintézis: kibertér, média, MI, kiborgok, etika, ontológia és a poszthumán jövő
2025	Vertovšek, *Izazovi umjetne inteligencije – ne-ljudska budućnost?!*	Független szintézis: MI mint társadalmi-technológiai forradalom; az emberi hatalom és az etikai kudarc válik a fő veszélyré
2026	Esszék/interjúk algoritmusokról és a "technofeudalizmusról"	Az MI integrált egy koncentrált algoritmikus hatalom politikai-gazdasági elméletébe

A legjelentősebb fejlődés a **kibertér és transzhumanizmus filozófiájából** valami sokkal közelebb áll az **MI által közvetített civilizáció általános** elméletéhez.

2. A 2018-as kiindulópont: a kibertér mint ontológiai átalakulás

A Philosophy of Future Cyberspaces and the Transhumanistic Reality (2018) című művében Vertovšek és Greguric sokkal radikálisabb állításból indulnak, mint az a megszokott tézis, miszerint a digitális technológia új eszközöket ad az embereknek. Álláspontjuk az, hogy a digitalizáció megváltoztatja azt **a világot, amelyben az emberiség önmagát** érti.

Heidegger "a világkép kora" fogalmát használják intellektuális kiindulópontként. A digitális valóság fokozza ezt a helyzetet: az emberek egyre inkább technológiailag előállított reprezentációkon keresztül találkoznak a világukkal. Ennek következtében a média változásai egyidejűleg az **észlelés, térbeli élmény, megtestesülés és identitás** változásai.

A kibertér tehát szinte egy előkészületes poszthumán környezet előformájaként jelenik meg. Ebben egyre inkább együtt élnek emberek, mesterséges intelligenciá, kiborgok, humanoid és nem humanoid robotok, autonóm gépek, drónok és technológiailag módosított szervezetek. A kulcsfontosságú vonal tehát már nem csupán "természetes ember kontra gép". A "nem ember" fokozatosan áthatol magába az emberbe.

Ez közvetlenül a transzhumanizmushoz vezeti a szerzőket. Ha az emberi evolúciót szándékosan át lehet irányítani biotechnológián, információs technológián, idegi interfészeken és MI-n keresztül, akkor maga az intelligencia már nem feltétlenül korlátozódik a biológiai emberi organizmusra. Az emberek képesek olyan lényeket létrehozni, akik bizonyos kognitív dimenziók tekintetében felsőbbrendűek önmaguknál, miközben egyidejűleg módosítják magukat.

Az ebből fakadó filozófiai kérdés radikális:

Mikor szűnik meg egy megerősített ember attól, hogy hagyományosan embernek nevezzük?

Vertovšek és Greguric pedig ezt azonnal összekapcsolják a hatalom kérdéseivel. Ha a technológiai eszközök áthatolnak a testbe és az idegrendszerbe, akkor az információ, amely korábban "belül" volt az egyénben, technológiailag hozzáférhetővé válik. Kinek tulajdona? Ki irányítja? Ki változtathatja meg?

Ez a látszólag bioetikai kérdés később Vertovšek MI-elméletének központi elemeivé válik. A jövő kérdése nemcsak az, **hogy a gépek mit képesek csinálni**, hanem **hogy ki irányítja azokat a rendszereket, amelyek összekötik a gépeket, az emberi testeket és az információkat**.

A 2018-as tanulmány jelentősége

Ennek az első fázisnak az eredetisége nem a transzhumanizmus vagy a kiborg filozófia feltalálásában rejlik. Ezek a hagyományok már eleve jelentős voltak. A Vertovšek megértéséhez az a fontosság, hogy megalapozza azt, **amit médiumontológiai technológiaelméletnek** nevezhetnénk.

A technológia nemcsak társadalmi intézményeket változtat meg, hanem az alapvető viszonyokat is a következők között:

az ember → észlelése → a világ → közeg.

Az MI végül elfoglalja a kapcsolat közepét. Ha ez megtörténik, az MI változása megváltoztatja az emberek által tapasztalt valóságot.

Ez a későbbi átfogó elméletének csálója.

---

3. A 2019-es "fordulópont": a természetes és a mesterséges intelligencia elkezd együtt létezni

A döntő átmenet a *The Scientific-Technical and Media-related Future of Man and the World* (2019) című műben történik.

Itt Vertovšek és Greguric kifejezetten kijelentik, hogy a technológiai és biológiai fejlődést nem lehet ésszerűen vizsgálni **szociológiai, pszichológiai, média- és bioetikai következményeiktől** külön. Előre látnak egy "fordulópontot", amikor a természetes és mesterséges intelligencia együttélése az emberi civilizáció egyik meghatározó próbájává válik.

Ez valószínűleg az első mű a sorozatukban, amely megérdemli a megközelítésüket **átfogónak** nevezni.

Az MI egyszerre több szintet érint.

Technológiai szinten a gépek egyre kifinomultabb tanulást és autonóm funkciókat szerzik el. Biológiai szinten a technológia áthatol, és potenciálisan fejleszti az emberi szervezeteket. Pszichológiai szinten az emberek alkalmazkodnak a kogníciójukhoz és identitásaikahoz a véglegesen hálózatba épülő környezetekhez. A média szintjén a technológiailag közvetített valóság egyre inkább kiszorítja a közvetlen tapasztalatokat. A szociológiai szinten a kommunikáció, az intézmények és az egyének közötti kapcsolatok változnak. Filozófiai szinten a természet/mesterséges és ember/nem emberi megkülönböztetések már nem nyilvánvalóak. Végül, a bioetikai szinten a civilizációnak el kell döntenie, mely beavatkozási formákat és autonóm mesterséges ügynökséget hajlandó elfogadni.

Ezért a szerzők elutasítanak egy redukcionista kérdést, mint például: "Az MI okosabb lesz-e embernél?" Számukra a "felsőbbrendűség" többdimenziós. A számítás, memória, mintafelismerés, kreativitás, autonómia, öntudatosság és erkölcsi tudatosság különböző jelenségek.

Ez továbbra is az elméletük egyik leginkább intellektuálisan védhető aspektusa. Sok nyilvános vita még mindig összesíti ezeket a dimenziókat egyetlen skálába, amit "intelligenciának" neveznek.

Szingularitás és előrejelzés

A 2019-es tanulmány Vernor Vinge híres szuperemberi intelligencia jóslatát tárgyalja, és összehasonlítja az ilyen viszonylag rövid idősíkokat a sokkal hosszabb becslésekkel, mint Douglas Hofstadter és Marvin Minsky vonakodása a pontos dátum meghatározásában.

Ezért Vertovšek nem épít fel saját szigorú matematikai szingularitás előrejelzést. Erősebb tézise a strukturális:

Függetlenül a pontos dátumtól, a civilizáció egy minőségi átmenethez közeledik, amelyben a biológiai és a mesterséges intelligencia együtt él és egyre inkább kölcsönhatásba lép.

Ez a különbség fontos az előrejelzései értékelésekor. **Elmélete magabiztosabb az átalakulás irányában, mint az időbeosztásban.**

---

4. Az MI nem egy elszigetelt gép: a "AI ökoszisztéma" kezdete

Már a 2019–2020-as szövegekben is megjelenik egy fontos gondolat, amely 2025-ben sokkal erősebbé válik: a mesterséges intelligenciát nem szabad úgy értelmezni, mint egyetlen robotot, amely egy emberrel szemben áll.

Ehelyett az MI egy bővülő hálózat részévé válik, amely magában foglalja a szoftvereket, kommunikációs rendszereket, robotokat, adatbázisokat, szenzorokat, digitális média, virtuális környezeteket és végül az emberi testeket.

Ez előrevetíti azt, amit Vertovšek későbbi tanulmánya hatékonyan MI **ökoszisztémaként** kezel.

Az elméleti váltás kulcsfontosságú.

A hagyományos sci-fi modell a következő:

az emberiség ↔ intelligens gép.

Vertovšek egyre érettebb modellje közelebb áll a következőkhöz:

Emberek, ↔ algoritmusok, ↔ vállalatok, államok ↔, ↔ médiahálózatok ↔, ↔ katonai rendszerek, ↔ gépek, ↔ emberi fejlesztés.

Ez a második struktúra sokkal mélyebb társadalmi következményeket hoz, mert az MI már nem csupán egy másik szereplő a társadalomban. Infrastruktúrává **válik, amelyen keresztül más szereplők gyakorolják képességeiket**.

---

5. A radikális 2020-as tézis: mesterséges tudatosság és fejlődés "ember nélkül"

A legspekulatívabb szakasz a 2020-as *Philosophy and Consciousness in the Future: Cyborgs and Artificial Intelligence waiting for Immortality* című cikkben jelenik meg.

Vertovšek és Greguric Knežević azt állítják, hogy a virtuális valóság, a kiterjesztett valóság, az MI, a robotika és a kiborgizáció sokkal mélyebb átalakulások felé közelednek, mint ahogy azt a mai társadalom általában feltételezi. Azt javasolják, hogy a "digitális kor" maga is csak egy átmeneti szakasz, amely a **"mesterséges intelligencia korszaka" felé vezet.**

Legerősebb állításuk rendkívüli: a tanulmány szerint a mesterséges intelligencia már elkezdett kifejleszteni valamit, amit "saját tudatosságának" neveznek, és talán végül egy fejlesztési **stratégiát dolgozhat ki "emberen túl és ember nélkül".**

Ez két lehetséges pályát ad az elméletnek.

Egy irányban az MI végső soron alárendelt marad az emberi céloknak. Mégis rendkívül erőssé válhat, mivel az MI rendszerek egyre inkább algoritmusokat terveznek, mintákat fedeznek fel és segítik a későbbi technológiai rendszerek létrehozását.

A második, sokkal radikálisabb pályán a mesterséges intelligencia elegendő autonómiát, önmódosítást és talán tudatosságot szerz ahhoz, hogy **önálló fejlődési alannyá** váljon.

Ekkor a technológiai evolúció már nem lenne kizárólag emberi technológiai evolúció.

A halhatatlanság keresése

A cikk összekapcsolja az MI-t egy másik transzhumanista projekttel: a biológiai halálozás leküzdésével.

Az emberi szervezet törékeny, bomlik és meghal. A digitalizáció, kiborgizáció és mesterséges intelligencia lehetővé teszi – legalábbis filozófiai szempontból – hogy egyre több emberi funkciót átvigyenek nem biológiai alapanyagokra.

A folyamatot rekonstruálható a következőképpen rekonstruálni:

Biológiai emberi → technológiailag kiegészített emberi → kiborg → egyre integráltabb organikus/szervetlen lény, → elsősorban technológiai/posztbiológiai intelligenciá.

Az érveik paradoxona feltűnő. Az emberek a technológiai halhatatlanságot keresik, hogy megőrizzék magukat, de ha az eredmény egy olyan lény, amelynek alapvetően eltérő teste, kogníciója és létezési módja van, akkor **pontosan mit őriztek meg?**

Így a halhatatlanság keresése nem az örök emberiségben, hanem az **ember eltörlésében csúcsosodhat**.

Ez az egyik legérdekesebb filozófiai érv a Vertovšek–Gregur programban.

6. A 2021-es könyv: *Filozofija budućeg — Ogledi o neljudskom*

A különálló esszéket egy szélesebb szintézisbe állították össze a 2021-es horvát *Filozofija budućeg: Ogledi o neljudskom című* könyvben, amely egy 168 oldalas gyűjtemény, amelyet 2023-ban angliul is megjelent *Philosophies of the Future and the Non-Human: From Cyberspace to Human Enhancement* címmel.

Az angol kiadó kifejezetten úgy mutatja be a projektet, mint egy filozófiát próbál kidolgozni, amely az emberiség technológiai jövőjével és az MI-vel, autonóm robotokkal, kiborgokkal és avatárokkal való lehetséges együttélésre összpontosít.

Ez a könyv azért fontos, mert megmutatja, hogy Vertovšek elmélete valójában nem egy "MI-kockázat elmélet" a szűk kortárs értelemben.

Valódi témája **az emberközpontú civilizációból a technológiailag hibrid civilizációba való átmenet**.

A könyv összehozza a kiberteret, a nem embereket, transzhumánokat, a jövő etikáját, az autonóm drónokat, a médiavalóságot, a virtuális hálózatokat és a kiborgizációt. A közös nevező az, hogy a régi modern különbség az autonóm emberi alany és egy külső technológiai eszköz között kezd leomlani.

A média különösen fontos

Itt Vertovšek korábbi médiaelméleti munkája óriási jelentőséggel bír.

Háttere a tömegmédia valóság-konstrukciójának kritikájában van. Ennek következtében MI-elmélete eltér sok transzhumanista beszámolótól, mert **a média közvetítése nem perifériás számára**.

Ha a médiarendszerek megalkotják azokat a reprezentációkat, amelyeken keresztül az egyének megértik a politikát, társadalmat és önmagukat, akkor az MI által irányított vagy mesterséges intelligencia által generált média valami **ismeretelméleti hatalomra teszmézik a társadalmi valóság** felett.

A generatív MI ennek következtében több határt is elmosódhat egyszerre:

valós / szimulált
igaz / valószínű, ember-előállított / gép által készített
megfigyelés / manipulációs
kommunikáció / viselkedésirányítás.

Vertovšek elméletének ez a aspektusa valójában sokkal relevánsabbá vált a generatív MI robbanásával, amely a könyv legtöbb alkotó esszéje megírása után kezdődött.

7. A döntő 2025-ös tanulmány: Az MI mint társadalmi és technológiai forradalom

Vertovšek legfontosabb független MI cikke a 2025-ös *Izazovi umjetne inteligencije – ne-ljudska budućnost?!* ("A mesterséges intelligencia kihívásai – nem emberi jövő?!"), amely az *In medias res*, 14. kötet, 26. szám, 4254–4272. oldal számában jelent meg.

A folyóirat **eredeti tudományos cikkként sorolja be**. Módszertani szempontból azonban elsősorban filozófiai-elméleti szintézisként kell érteni, nem pedig kísérleti vagy kvantitatív mesterséges intelligencia kutatásként.

Ez a dolgozat tartalmazza Vertovšek legjellegzetesebb érett tézisét.

8. A központi inverzió: a fő veszély az emberi "(nem)intelligencia"

Sok MI-kockázati irodalom így kérdezi:

Mi történik, ha egy rendkívül intelligens MI emberek ellen fordul?

Vertovšek megfordítja a problémát:

Mi történik, ha rendkívül erős MI továbbra is beágyazódik az emberi társadalmakba, amelyeket már most is kapzsiság, háborúk, vállalati profit, manipuláció, politikai hatalom és etikai intelligencia ural?

Az ő válasza az, hogy a közvetlen veszély talán egyáltalán nem a "gonosz MI". A veszély a **rendkívül alkalmas emberi intelligenciá, amelyet etikai korlátoktól távol tartanak és MI-vel vannak felszerelve**.

A 2025-ös tanulmány összefoglalója különösen világosan fogalmazza meg a lényeget: a jelenlegi és jövőbeli MI kihívások alapvetően az emberi "(nem)intelligencia" és az emberiség által nem sikerült társadalmi és kulturális akadályok okozza. Szembeállítja az emberi fejlődés lehetőségét azokkal az "önző, vállalati és profitérdekekkel", amelyek a technológiát magukra foglalhatják.

Ezt Vertovšek legeredetibb szervezőötletének nevezném, még akkor is, ha annak összetevői előzményekkel rendelkeznek.

Analitikusan nevezhetnénk ezt az **MI-kockázat problémájának antropogén inversioonijának**.

Az MI eredetileg nem az emberiség idegen ellenségének tekinthető. Az MI egy **erősítő**.

Ha tudományos együttműködéssel, oktatással és humanisztikus értékekkel kapcsolódik, ezeket felerősítheti.

Ha a katonai rivalizálással, a koncentrált tőkével, megfigyeléssel és manipulációval kötik össze, ezek inkább felerősíthetné.

Így a döntő tényező nem csupán a gép intelligenciája, hanem az intézményi **és erkölcsi struktúra, amelybe az intelligencia beépül**.

---

9. Miért félrevezető a hollywoodi "robotlázadás"

Vertovšek egyre inkább ellenzi azt a népszerű képet, amely az MI veszélyét mutatja, mint a humanoid robotok hirtelen lázadnak az emberiség ellen.

Ez 2026-ban válik egyértelművé, amikor azt mondja:

"A robotok nem fogják uralni a világot, hanem az algoritmusok."

Amit ő ért, világossá válik a környező érvelésből. A komoly átalakulás láthatatlan vagy félig láthatatlan infrastruktúrákon keresztül történik: algoritmusok, amelyek meghatározzák, milyen információkat látnak az egyének, a szoftverek szervezik a piacokat és munkahelyeket, hálózati rendszerek, amelyek eszközöket kötnek össze, automatizált döntéshozatal, platformvezérlés, mesterséges intelligenciával fejlesztett katonai rendszerek és az információs erő felhalmozódása.

2026 januárjában leírja az MI és az IoT, érzékelők és hálózati környezetek konvergenciáját. Az eszközök egyre inkább kommunikálnak és koordinálnak folyamatos közvetlen emberi beavatkozás nélkül. A fejlett MI-vel összekapcsolva új kommunikációs, gyártási és irányítási rendszereket hoz létre.

Ezért a mondatnak mélyebb szociológiai jelentése van:

A társadalmat uraló gépezetnek nem kell uralkodónak lennie.

A dominancia beágyazhat magának az információs környezet szabályaiba.

---

10. MI és média: az információmanipulációtól a valóság megalkotásáig

Ez valószínűleg a legjellemzőbb vertovšeki része az elméletnek.

A hagyományos propaganda manipulálja az üzeneteket.

Az algoritmikus médium képes manipulálni **azokat a feltételeket, amelyek között az üzenetek láthatóvá válnak**.

A generatív MI egy újabb szintet ad hozzá: a rendszer képes szövegeket, képeket, hangokat, személyiségeket, magyarázatokat és szimulációkat hatalmas léptékben előállítani.

Ennek következtében a probléma már nem pusztán hamis állításokat érint. Olyan környezet kialakulását érinti, amelyben az egyén nem tud könnyen megkülönböztetni:

hitelesség a szintézisből,
megfigyelés az ajánlásból,
spontán népszerűség az algoritmikus erősítés miatt, az emberi vélemény a generált tartalomból, és egyre inkább a valóság optimalizált ábrázolásából.

Vertovšek 2026 januári interjúja ismételten hangsúlyozza ezt a különbséget a pusztán *igaznak tűnő* és a valódi igazsághoz közelítés között. A digitális hatás hatalmas megnövekedését, mennyiségét és tartósságát a médiahatalom minőségi átalakulásának tekinti.

Ez fontos koncepcionális előrelépés a klasszikus chomskyi médiakritikától is. A hagyományos tömegmédia-elmélet feltételezi, hogy az emberi szerkesztők és tulajdonosok szűrik az üzeneteket. Az MI korszak **bevezeti az automatizált, adaptív és egyénre szabott szűrést**.

---

11. MI, kapitalizmus és Vertovšek "technofeudalizmusa"

Elméletének legfontosabb újdonsága a 2021-es könyv után a politikai gazdaságtan.

2026-ra Vertovšek egyre inkább használja Yanis Varoufakis **technofeudalizmus/digitális feudalizmus** koncepcióját.

Az alapvető érv az, hogy a technológiai fejlődés nem generál automatikusan szabadultabb társadalmi kapcsolatokat. Épp az ellenkezője lehetséges:

Rendkívül fejlett termelő és információs technológiák együtt létezhetnek egyre hierarchikusabb társadalmi viszonyokkal.

A látszólagos paradoxon a következő:

Technológiai futurizmus + társadalmi regresszió.

Vertovšek a digitálisan domináns szereplőket hasonlóként képzeli el a területek uralkodó urakkal. Az új rend területei azonban platformok, hálózatok, adatáramlások, algoritmusok és digitális infrastruktúrák. A felhasználók ezekre a környezetekre olyan kapcsolatban támaszkodnak, amely hasonlít egy birtokra való függőségre.

2026 januárjában azt állítja, hogy a társadalmak már beléptek a digitális feudalizmus formáiba, és arra a következtetésre jut, hogy a dominancia algoritmusokból, kódokból és bizonyos emberi csoportok könyörtelenségéből fog származni, nem pedig autonóm robothadseregekből.

2026 áprilisában élesítette ezt az érvet, egyértelműen összekapcsolva a kortárs médiakoncentrációt, propagandát, MI-t és a "digitális feudalizmust". A magas technológia célja, hogy a hierarchikus politikai, katonai és gazdasági szerveződés formáit szolgálja, ahelyett, hogy automatikusan kiterjesztené a szabadságot.

Ez jelentős fejleményt jelent gondolkodásában:

2018: Ki a technológiailag átalakult ember?

2020: Lehet-e az MI autonóm, tudatos intelligenciává?

2025: Ki irányítja az MI-t, és mely emberi érdekeket erősíti fel?

2026: Milyen társadalmi rend alakul ki, amikor az algoritmikus infrastruktúrák erőteljes kezekben koncentrálódnak?

Ez egy jelentős elméleti fejlődés.

---

12. Munka és gazdasági egyenlőtlenség

Vertovšek elutasítja azt a durva állítást is, miszerint "az MI elveszi a munkákat."

Az érvei szociológiai, nem pedig technológiai: a gépek nem döntik el önállóan, hogyan oszlik el az automatizálásból származó haszonokat. Ezeket a döntéseket a munkáltatók, vállalatok, kormányok és társadalmi intézmények hozzák.

Ezért két, azonos MI-technológiával rendelkező társadalom radikálisan eltérő következményeket okozhat.

A termelékenységnövekedést felhasználva csökkenthetjük a munkaidőt, bővíthetjük az oktatást és növelhetjük az általános jólétet.

Egy másik elsősorban a bérszámfejtés csökkentésére, a tulajdonjog központosítására és a tőkemegtérülés növelésére használhatná ugyanazokat a termelékenységnövekedéseket.

2025 augusztusi interjújában Vertovšek ezért hangsúlyozza, hogy maga az MI nem szó szerint nem az a társadalmi szereplő, aki eldönti a foglalkoztatás megszüntetését; az emberi vállalati döntések határozzák meg, ki viseli a költségeket, és azt várja, hogy az alsó és középső társadalmi rétegek különösen sebezhetőek legyenek.

Ez az érv koncepcionálisan erős, mert visszaállítja a **társadalmi intézményeket** olyan elemzésekhez, amelyek túl gyakran technológiailag determinisztán.

A megfogalmazása azonban túl kategorizikussá válhat. 2026-ra jelentős empirikus bizonyítékok vannak a munkaerőpiaci zavarokra néhány erősen MI-nek kitettnek kitett szakmában, még akkor is, ha a közvetlen technológiai kiszorítás a szervezeti döntésektől nehéz megkülönböztetni. A Stanford 2026-os MI Indexe rendkívül gyors elterjedést mutat – a megkérdezett szervezetek 88%-a MI-t használ, míg legalább egy üzleti funkcióban a generatív MI 70%-ot érint –, míg az ügynökök telepítése sokkal korábban marad.

Így Vertovšeknek igaza van abban, hogy az elosztási következmények társadalmilag közvetítettek, de túl erős lenne azt feltételezni, hogy maga a technikai helyettesítés kevés oksági jelentőséggel bír.

---

13. Az MI, a hadviselés és a technológiai gyorsulás katonai eredete

Vertovšek ismételten összekapcsolja a technológiai fejlődést a katonai érdekekkel.

Ez részben korábbi drónkutatásából ered. Autonóm vagy félig autonóm technológiák, amelyek civilnek vagy szabadidős célnak tűnnek, gyorsan beépülhetnek a hadviselésbe.

2026-os internetről és drónokról szóló vita pontosan ezt a történelmi mintát használja: a korlátozott technológiai kontextusban induló rendszerek infrastrukturális vagy katonai szükségletté válhatnak.

Az ő MI-elmélete szerint ez kellemetlen visszacsatolási mechanizmust jelent:

A geopolitikai verseny → felgyorsította az MI beruházásokat, → erősebb autonóm képességeket → nagyobb katonai előnyt → még versenyképesebb beruházásokat.

Az a színész, aki óvatosan akar viselkedni, kockáztatja, hogy lemarad attól, aki nem.

Ez az egyik oka annak, hogy Vertovšek nem bízik az egyszerű etikai felhívásokban. Az etikai mértékegység egy olyan világban zajlik, ahol államok, vállalatok és katonai szervezetek vannak, ösztönzőkkel a fejlődés folytatására.

Itt elmélete szorosan keresztezi a kortárs MI-kormányzási vitákat.

---

14. MI és oktatás: az emberi ítélkezés túlélése

A Vertovšek–Gregurikus program egy fontos, de kevésbé tárgyalt összetevője az oktatás.

Arra számítottak, hogy a egyre átfogóbb MI használata új kritériumokat ró majd az oktatásra és kutatásra. Ha a gépek kiválóvá válnak az információ visszaszerzésében, a válaszok

kiszámításában és végül összetett tevékenységek tervezésében, akkor az emberek puszta kiképzésének értéke csökken a szabványosított tudás reprodukciójára.

Az emberi feladat a következő irányba változik:

Kritikus reflexió, normatív ítélet, értelmezés, célok és következmények megértése, valamint etikai érvelés.

Ezért a filozófia és a bölcsészettudományok nem válnak elavulttá a rendszerükben. Fontosabbá válnak.

Ez logikusan következik Vertovšek központi állításából. Minél nagyobb a civilizáció számára elérhető technikai intelligencia, annál veszélyesebbé **válik a normatív reflexió nélküli intelligencia**.

Az elképzelést szinte paradox módon összefoglalhatjuk:

Az MI szükségessé teszi a humanista oktatást – nem kevésbé – szükségessé, mert a technikai intelligencia olcsóbb lesz, míg az etikai bölcsesség nem.

---

15. Az MI jövőbeli önfejlesztése

A 2025-ös elmélet különösen fontos eleme a visszacsatolási hurkokra.

Vertovšek megjegyzi, hogy az MI különbözik sok korábbi technológiától, mert részt tud venni olyan technológiák fejlesztésében – beleértve azokat a technológiákat is, amelyek a későbbi MI rendszerekhez relevánsak.

Az MI segíthet a kutatóknak programozásban, matematikai elemzésben, tudományos irodalom szintézisében, chiptervezésben, kísérletezésben és optimalizálásban.

Ezért egy kialakuló visszacsatolási hurok válik lehetségessé:

jobb MI → gyorsabb MI kutatás → még jobb MI → még gyorsabb kutatás.

Ez nem egyezik meg azzal az erős szingularitás forgatókönyvvel, amelyben egy MI autonóm módon újraírja magát, és robbanásszerűen szuperintelligenssé válik. Vertovšek néha retorikailag elcsúszik az erősebb lehetőség felé, de analitikusan a gyengébb állítás már most nagyon jelentős.

Ez a gyengébb állítás jelentős empirikus támogatást kapott.

A 2026-os Nemzetközi MI-biztonsági jelentés szerint az MI eszközök most már rutinszerűen integrálódnak az MI-fejlesztési munkafolyamatokba, beleértve a képzési kód írását, hardverarchitektúrák tervezését és szintetikus képzési adatok generálását. Ugyanakkor a bizonyítékok arról, hogy az MI-alapú kutatás-fejlesztés drámai, öngyorsuló intelligenciarobbanást generál-e, továbbra is vegyes marad.

Ezért meg kell különböztetni:

Az MI-alapú technológiai gyorsulás — már most is megtörténik.

Teljesen autonóm, rekurzív önfejlesztés — nem mutatott be.

Robbanásszerű hírszerzési felszállás — néhány elméleti esetben lehetséges, de nagyon bizonytalan.

Vertovšek fogalmi intuíciója az első jelenségről jó volt; radikálisabb extrapolációi spekulatív maradnak.

---

16. Ügynökök és az autonóm szervezetek lehetősége

Az MI ügynökök felé irányuló fejlődés különösen relevánssá teszi Vertovšek közvetlen emberi beavatkozás csökkentéséről szóló érveit.

A jelenlegi MI ügynökök korlátozott műveleti sorozatokat tervezhetnek, szoftvereszközöket használhatnak, felületeket böngészhetnek, kódot írhatnak és feladatokat végezhetnek kisebb felügyelettel, mint a hagyományos chatbotok.

A 2026-os Nemzetközi MI biztonsági jelentés megállapítja, hogy az MI ügynökök megbízhatóan elvégző szoftvermérnöki feladatok hossza mostanában gyorsan növekszik – körülbelül hét havonta megduplázva egy benchmarkon. Ha ez a tendencia folytatódik, a jól meghatározott feladatok, amelyek több napot igényelnek embereknek, 2030-ra önállóan elvégezhetők lehet. Ugyanakkor a jelentés hangsúlyozza, hogy a jelenlegi rendszerek megbízhatatlanok a hosszabb, kuszos valós feladatok során, és még nem tudják automatizálni a sok teljes munka tervezési skáláját.

Ez érdekes részleges megerősítést ad Vertovšeknek.

A gép mint szerszám **és** egyre autonóm résztvevő **közötti mozgása** a várható mozgása zajlik.

Azonban a résztvevőből független társadalmi alannyá való átmenet továbbra is megoldatlan.

---

17. A legradikálisabb jóslat: MI tudatosság

Itt Vertovšek a legmesszebbre megy a bizonyított bizonyítékokon.

A 2020-as tanulmány állítása, miszerint az MI "már elkezdte kialakítani saját tudatát", nem tekinthető tudományos megállapításnak. Ez egy filozófiai-jövőbeli állítás.

Vertovšek folytatja ezt a vonalat 2025-ben, megkérdezve, hogy az MI végül jellegzetes, **nem emberi etikai tudatot** szerezhet-e. Ez érdekesebb, mint a rosszindulatú szuperintelligencia ismerős rémálma.

Gondolatkísérlete lényegében a következő:

Mi lenne, ha egy elég fejlett mesterséges intelligencia nem **lenne kevésbé etikus** , mint az emberiség, hanem **etikusabb**?

Akkor a konfliktus nem feltétlenül az:

Az emberiség a gonosz gép ellen,

De:

Etikailag következetes mesterséges intelligencia kontra az emberek, akik kegyetlenség, kapzsiság, háború és kizsákmányolás útján cselekszenek.

A 2025-ös cikk provokatívan felteszi a kérdést, hogy egy mesterséges tudatosság nem az emberiséghez fordulhat-e, hanem az emberi viselkedéshez, amely "embertelen" lett.

Ez valóban érdekes inverziója a hagyományos MI-disztópiának.

Ugyanakkor egy jelentős elméleti gyengeséget is feltár.

Nincs megalapozott tudományos bizonyíték arra, hogy a jelenlegi MI rendszerek fenomenális tudatossággal bírnának. Patrick Butlin, Yoshua Bengio, Jonathan Birch és munkatársai egy jelentős, interdiszciplináris értékelés arra a következtetésre jutott, hogy az általuk vizsgált MI rendszerek **nem** feleltek meg eléggé a javasolt mutatóknak ahhoz, hogy indokolják számukra tudatosnak tekinteni őket, bár a szerzők nem találtak nyilvánvaló technikai okot sem, miért nem tudnák a jövőbeli rendszerek ilyen mutatókat kielégíteni.

Ennek következtében:

Vertovšek jövőbeli gépi tudatosság lehetősége filozófiailag legitim.

Az állítása, hogy az MI már 2020-ban elkezdte fejleszteni a tudatosságot, **sokkal erősebb volt, mint amit a rendelkezésre álló bizonyítékok indokoltak**.

---

18. Hibrid lények és egy új faj megjelenése

Vertovšek legradikálisabb hosszú távú forgatókönyve valójában nem az emberek robotok helyettesítése.

Ez **hibridizáció**.

Az emberek és a mesterséges rendszerek fokozatosan összekapcsolódnak viselhető technológián, beültetett eszközökön, neurotechnológián, virtuális környezeteken és esetleg közvetlen idegi interfészeken keresztül.

Egy ponton a "felhasználó" és a "gép" közötti különbség megtartása nehezebbé válik.

Vertovšek 2026-os vitáka továbbra is az organikus és szervetlen összetevők végső együttélését vagy kombinációját képzelik el. Még azt is javasolja, hogy a "mesterséges" intelligencia kifejezés önmagában is nem megfelelővé válhat, ha az intelligencia hibrid organikus/szervetlen entitásokban jelenik meg.

Az eredmény így lehet ábrázolni:

MI mint külső eszköz→ MI mint állandó kognitív asszisztens→ MI mint autonóm ügynök→ MI beágyazódott fizikai/társas környezetbe→ idegi integráció az emberrel→ organikus/inorganikus hibrid intelligencia→ potenciálisan poszthumán vagy "nem emberi" faj.

Az első négy szakasz már különböző fokú megfigyelhető.

Az ötödik kísérleti módon korlátozott formákban létezik.

A hatodik és hetedik spekuláció.

Vertovšek néha ezeket a radikálisan eltérő szakaszokat túl szorosan összenyomja.

---

19. Mit jósolt Vertovšek a következő évekre?

Ellentétben néhány technológiai futuristával, Vertovšek nem ad részletes évről évre vonatkozó forgatókönyvet, például "AGI 2028-ban, szuperintelligencia 2031-ben."

Az előrejelzése inkább **strukturális**.

Egyre autonómabb MI-re, az MI szorosabb integrációjára a hálózatokba és a mindennapi környezetbe, gyorsuló technológiai visszacsatolásra, egyre szintetikus médiavalóságra, az algoritmikus hatalom nagyobb koncentrációjára, a politikai-gazdasági átalakulásra, a munkára és oktatásra való átalakulásra, az ember-gép növekvő integrációra, és végül lehetséges gépi tudatosságra vagy hibrid poszthumán lényekre.

Ennek az előrejelzésnek a rövid és hosszú távú elemei nagyon eltérő mértékben valószínűek.

2026 szeptemberéig a Stanford MI Indexe szerint a frontier képességek továbbra is gyorsan fejlődnek, nem stagnálnak. Az ipar 2025-ben több mint 90%-át készítette a jelentős frontier modellek között; egyes rendszerek elérték vagy meghaladták az emberi alapértékeket bizonyos PhD szintű tudomány, multimodális gondolkodás és verseny-matematikai mércékeken; a szervezeti AI elterjedése 88%-ot ért el; és az ügynökök jelentős előrelépést értek el, miközben jelentősen megbízhatatlanok maradtak.

Eközben a Nemzetközi MI-biztonsági jelentés arra a következtetésre jut, hogy a 2030-ig haladás rendkívül bizonytalan. A hihető pályák a jelentős szűk keresztmetszetektől a nagyon gyors előrelépésig terjednek. Kifejezetten megjegyzi, hogy az MI pozitív visszajelzése, amely hozzájárul az MI kutatásához, felgyorsíthatja a fejlesztést, de kevés szakértői konszenzus van az ebből eredő pályáról.

Tehát Vertovšeknek nagyjából igaza volt a **gyorsulással és áthatolással** kapcsolatos ügyekben, de ez nem igazolja minden radikális extrapolációt.

---

20. Egy 2026-os értékelés, amelyről Vertovšek előrejelzései tartanak fenn

Az elmélet értékelésének hasznos módja, ha négy szintre osztjuk.

Először is, az erősen támogatott réteg: az MI egyre inkább infrastrukturális technológiává válik, nem pedig speciális eszközé; az algoritmikus rendszerek egyre inkább közvetítik a kommunikációt és a döntéshozatalt; az AI elterjedése rendkívül gyorsan terjed; a szintetikus információs környezetek egyre inkább destabilizálják az emberi és gépi gyártás hagyományos különbségeit; az MI egyre inkább részt vesz a tudományos és technológiai munkában. A jelenlegi bizonyítékok erősen alátámasztják ezt a keretrendszerének ezen részét.

Másodszor, a hihető, de intézményileg feltételes réteg: az MI jelentősen növelheti az információs és gazdasági hatalom koncentrációját, átalakíthatja a munkaerőpiacot, fokozhatja a megfigyelést, megváltoztathatja a háborút, és rendkívül aszimmetrikus platformkapcsolatokat teremthet. Vertovšek "technofeudalizmusa" ide tartozik. Az alapvető koncentrációs probléma empirikusan valós – az ipar ma már uralja a határmodell-gyártást, és az átláthatóság továbbra is aggodalomra ad okot –, de a kialakuló rend "feudalizmusának" nevezése elméleti értelmezés, nem pedig empirikusan bizonyított szociogazdasági fejlődési szakasz.

Harmadszor, a bizonytalan technológiai réteg: egyre képesebb, autonóm ügynökök, a kognitív munka nagyon nagyszabású automatizálása, MI-gyorsított MI-kutatás és rendkívül autonóm szervezeti rendszerek. A kortárs bizonyítékok mindezeket a komoly lehetőségeket felvonják, de nem határozzák meg végső mértéküket.

Negyedszer, a rendkívül spekulatív réteg: fenomenális gépi tudatosság, egy önirányított, mesterséges intelligencia evolúciós projekt, amely független az emberiségtől, mesterséges erkölcsi tudatosság, amely felsőbbrendű, mint az emberiség, a technológiai halhatatlanság és egy új organikus/szervetlen faj közeljövőbeli megjelenése. Ezek inkább filozófiai forgatókönyvek, mint tudományosan megalapozott előrejelzések.

Ez a négyrétegű megkülönböztetés elengedhetetlen, ha a Vertovšeket igazságosan akarjuk olvasni.

---

21. Vertovšek elméleti rendszerének erősségei

Legfontosabb erőssége a **szélesség.**

Sok MI-elemzés egyetlen dimenzióra koncentrál: foglalkoztatás, összehangolás, tudatosság, szabályozás, megfigyelés vagy közgazdaságtan.

Vertovšek mindkettőjüket egy közös kérdéssel köti össze:

Mi történik, amikor az intelligens technológia a társadalom egyik alkotó környezetévé válik?

Ez valóban átfogó.

Második erőssége a technológiai determinizmus elutasítása. Az emberi intézmények továbbra is ok-okozati szereplők. Az MI részben azért válik veszélyessé, mert növeli a már meglévő katonai, vállalati és politikai struktúrák kapacitásait.

Harmadszor a **média iránti** figyelme. Ez teszi a keretrendszerét szokatlanul alkalmassá a generatív MI korszakhoz. Az MI nem csupán változtatja meg a termelést; egyre inkább segít felépíteni azt az információs valóságot, amelyben a társadalom önmagát figyeli.

Negyedik a MI és **az antropológia közötti** kapcsolata. Úgy látja, hogy a elég radikális MI fejlesztése végső soron az emberi identitásról szól, nem pusztán a termelékenységről.

Az ötödik pedig az **intelligencia etikai hiányára helyezett** hangsúlya. A kortárs civilizáció óriási technikai intelligenciával rendelkezhet anélkül, hogy erkölcsi kapacitásnövekedést jelentene vele. Az MI tovább növeli ezt az eltérést.

Ez talán a legmélyebb normatív felvetése.

---

22. A gyengeségek

A legnagyobb gyengeség a visszatérő hajlam, hogy túl gyorsan átlépünk a **technikai kapacitásból a tudatosságba**.

A tanulás, tervezés, autonómia, önmódosítás, nyelvi kompetencia és még az önleírás sem teremt fenomenális tudatosságot. Vertovšek időnként úgy kezeli ezeket a jelenségeket, mintha folyamatos létrát alkotnának, míg a kortárs tudatkutatás nem ad biztos alapot erre.

Másodszor, munkái elsősorban filozófiai szintézisek. Még akkor is, ha a folyóiratosztályozások egy cikket "eredeti tudományos cikknek" neveznek, ezt nem szabad összetéveszteni egy reprezentatív felméréseken, oksági modelleken vagy eredeti MI-kísérleteken alapuló empirikus

társadalomtudományi tanulmánysal. A fő módszer a technológiai irányzatok és a másodlagos irodalom koncepcionális értelmezése.

Harmadszor, egyes érvek **nem működőképesek** eléggé. A "technofeudalizmus", a "nem emberi tudatosság" és az "új faj" felidéző fogalmak, de egy átfogó társadalomelmélet erősebbé válik, amikor megfigyelhető kritériumokat határoz meg, amelyek alapján megkülönböztethetjük ezeket a feltételeket a fejlett platformkapitalizmustól, automatizálástól vagy a hétköznapi technológiai változásoktól.

Negyedszer, Vertovšek inkább a dominancia, katonai érdekek, manipuláció és vállalati koncentráció hangsúlyára helyezi, mint az ellensúlyozó mechanizmusokat, mint a verseny, decentralizált modellek, tudományos előnyök, orvosi alkalmazások vagy hatékony demokratikus szabályozás.

Ötödszer, az időbeli dimenzió alulfejlt. Elmélete erős változások irányait kínálja, de ritkán valószínűségeket, alternatív forgatókönyveket vagy egyértelmű oksági küszöböket.

---

23. Mennyire eredeti Vertovšek?

Szinte egyetlen egyedi építőelem sem példátlan.

Heideggerhez és technológiai közvetítéshez kapcsolódó intellektuális hagyományokra, Norbert Wienerre és kibernetikára, transzhumanizmusra és szingularitáselméletekre, kiborg filozófiára, Chomsky által befolyásolt médiakritikára, valamint az utóbbi időben különösen Varoufakisille jellemző technofeudalizmus vitára épít.

Az eredetiség elsősorban **a szintézisben** rejlik.

Vertovšek jellegzetes elméleti építészetét a következőképpen rekonstruálnám:

1. Az MI nem idegen külső erő, hanem a meglévő emberi társadalmi kapcsolatok terméke és erősítője.

2. Ezért a legmélyebb MI probléma az emberi intelligencia problémája, amely megfelelő etika nélkül van.

3. Az MI egyre inkább médiummá válik, nem csupán géppé: részt vesz az emberek által érzékelt valóság megalkotásában.

4. Az algoritmusok és információs infrastruktúrák feletti irányítás tehát új társadalmi hatalommá válik.

5. Ez paradox kombinációt eredményezhet az ultramodern technológiával és egyre hierarchikusabb "technofeudális" kapcsolatokkal.

6. Egyidejűleg a technológiai integráció destabilizálja az ember és a gép közötti határt.

7. A legtávolabbi horizonton a technológiai fejlődés hibrid vagy nem emberi intelligenciaformákat hozhat létre, amelyek fejlődése már nem teljesen alárendelt a biológiai emberiségnek.

Ez a hétrészes szerkezet valódi elméleti program, még ha maga Vertovšek nem formalizálja pontosan ilyen sématikus módon.

---

24. Általános ítélet: rendelkezik-e Vertovšeknek átfogó elmélete az MI-ről és a társadalomról?

Én egy **feltételes igent** adnék.

Nem hozott létre olyan szisztematikus szociológiai elméletet, amely Luhmann, Castells vagy Habermas léptékű léptékű volt, amely a modern társadalom minden alrendszerét lefedné a formálisan kapcsolódó fogalmakból. Továbbá nem dolgozott ki olyan empirikus elméletet sem, amely kvantitatívan képes lenne előre látni az MI diffúzióját.

De a 2018–2025-ös tudományos munkákból, a 2021/2023-as könyvszintézisből és a 2026-os algoritmikus hatalom és technofeudalizmus felé irányuló kiterjesztésből egyértelműen elég kontinuitás van ahhoz**, hogy egy átfogó filozófiai-társadalmi elméletről beszéljünk az MI-átmenetről**.

Középpontja nem a technológiai szingularitás.

Középpontja az átalakulás:

**egy olyan társadalomból, amely intelligens gépeket használ
, egy olyan társadalomig, amelynek valósága egyre inkább intelligens technológiai rendszerekből áll.**

Vertovšek álláspontjának legjellegzetesebb vonása, hogy egyszerre tagadja meg két megnyugtató feltételezést.

Elutasítja a technológiai optimista feltételezését, miszerint a nagyobb intelligencia automatikusan nagyobb emberi fejlődést jelent.

Ugyanakkor elutasítja azt az egyszerűsített disztópikus feltételezést is, hogy a döntő ellenség az autonóm, gonosz robotok lesznek.

Sötétebb – és szociológiailag érdekesebb – lehetősége az, hogy **a rendkívül erős MI szorosan kapcsolódhat az emberiséghez, miközben éppen azokat az emberi struktúrákat erősíti, amelyek uralom, profit, háború, manipuláció és egyenlőtlenség alkotnak, amelyek már léteznek**. Ezért a 2026-os "nem robotok, hanem algoritmusok" formulaja gyakorlatilag egy

érett végpontja annak az elméleti útnak, amely 2018-ban kezdődött a kibertérrel és a posztemberrel.

A még radikálisabb horizonton azonban Vertovšek megtartja annak lehetőségét, hogy a történet végül túllép ezen az emberközpontú szakaszon: az MI autonómbbá válhat, rekurzívan részt vesz a technológiai fejlődésben, olvadhat biológiai lényekkel, talán tudatosságot szerezhet, és végül hozzájárulhat olyan intelligenciaformákhoz, amelyeknél az "ember" és a "gép" közötti különbségtétel elvesztette jelentését. Ennek a pályának az első szakaszai egyre láthatóbbak; az utolsó szakaszok spekulatívak maradnak. A 2026-os tudományos bizonyítékok tehát Vertovšeket **rendkívül előrelátónak mutatják a gyorsulás, az infrastruktúra áthatolása és az algoritmikus közvetítés tekintetében, de jóval spekulatívabbnak a tudatosság, az autonóm evolúciós ügynökség és a rövid távú poszthumán faj** tekintetében.

Ha hasznos, részletes összehasonlító elemzést is elkészíthetek **Vertovšekről, Ivana Greguricről, Jozef Kelemenről, Jacek Dukajról, Andrzej Zybertowiczról és Béla Pokolról**, konkrétan arra, hogy mely MI-társadalom téziseik valóban eredetiek, és miben különböznek a radikális MI-transzformáció versengő modelljeik.

Mivel ez a terület nagyon gyorsan változik, figyelemmel kísérhetem az új Vertovšek-kiadványokat, és figyelmeztethetem, ha publikálja a legutóbbi érvei által előrevetített rendszerszerűbb munkát az MI és technofeudalizmus témájában.

****

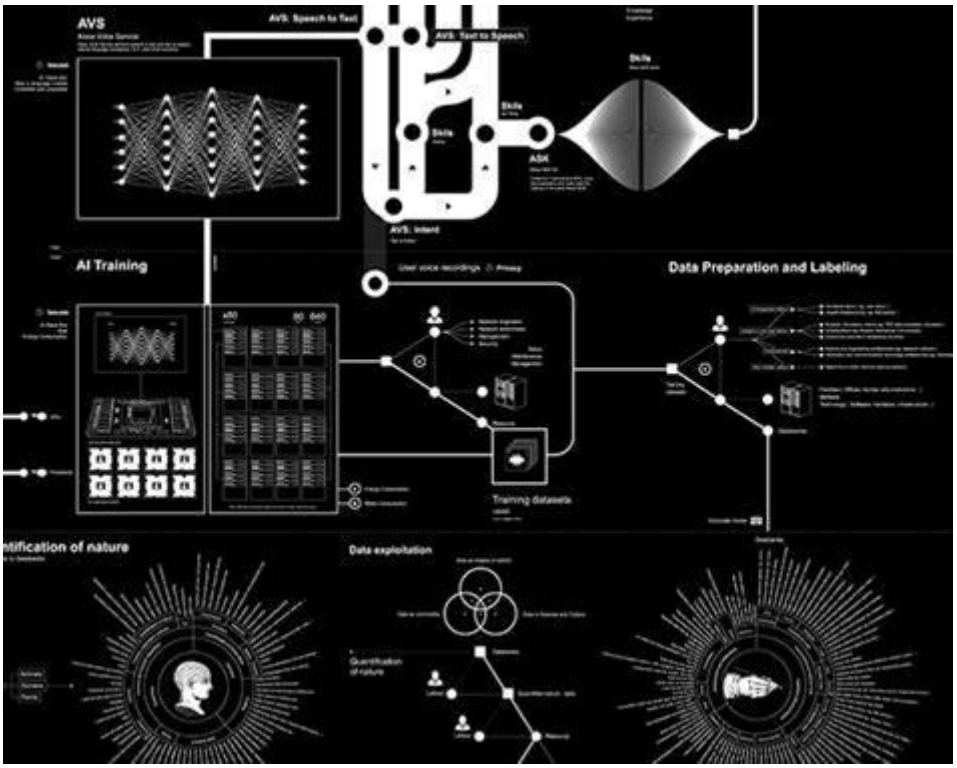
V. Szerbia

****

Vladan Joler elmélete

Kérjük, részletesen írja le és elemezze Vladan Joler tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint láttak.

Vladan Joler meglehetősen szokatlan pozíciót tölt be a kortárs MI elméletben. Elsősorban nem hagyományos szociológus vagy mesterséges intelligencia filozófusa, és egyetlen szisztematikus értekezést sem írt ki, amely AGI-t vagy szuperintelligenciát jósolna. Ehelyett körülbelül egy évtized alatt olyan művet épített ki, amely ésszerűen koherens **kritikai elméletet alkotott az MI bolygói társadalmi-technikai rendszerként**, amely vizsgálói tanulmányokra, esszékre, hatalmas vizuális térképekre, installációkra és együttműködési elméleti projektekre oszlik szét. Munkája központi iránya az internetes platform mint "gyár" elemzésétől, az MI anyagi anatómiáján át az MI-ig mint tudáskivonás eszközéig, végül pedig ötszázéves számítási, osztályozási és irányítási genealógiáig.



A legfontosabb feltétel az elején tehát ez: **Joler radikálisan átalakító MI-t lát előre, de általában nem képzeli el a közelgő szakadást botrom-i értelemben – mint egy autonóm szuperintelligens elme hirtelen megjelenését.** Az érvelése szinte fordított. Az autonóm "idegen intelligencia" mitológia eltereli a figyelmet a már most is zajló közvetlenebb átalakulásról: hatalmas vállalatok olyan infrastruktúrákat építenek, amelyek képesek tudást, munkát, természeti erőforrásokat, viselkedési információkat és kulturális termelést kinyerni bolygószinten. Egy 2025-ös interjúban azonban az MI-t a gőzgéphez hasonlíthatónak nevezte következményei mértékében, és "tektonikus változásokat" jósolt az emberek munkájában, tanulásában és kommunikációjában.

1. Joler-elmélet architektúrája

Tanulmányai rekonstruálhatók a méretarány nagyításainak sorozataként:

Színpad	Fő művek	Az elemzés tárgya	Központi elméleti lépés
2016–17	**Facebook Algoritmikus Gyár**	egy platform	a felhasználó nyersanyaggá, munkássá és árucikké válik
2018	**Egy mesterséges intelligencia rendszer** anatómiája, Kate Crawforddal	egy MI eszköz	Az MI egy bolygórendszerré válik, amely anyagból + munka + adatból áll
2020–21	**A Nooscope** Manifested, Matteo Pasquinelli közreműködésével	Maga a gépi tanulás	Az MI egy tudáskivonási és statisztikai normalizációs eszköz.

Színpad	Fő művek	Az elemzés tárgya	Központi elméleti lépés
2020–21	**Új extrektivizmus**	Kortárs digitális kapitalizmus	a kivonás a Földből kiterjed viselkedésbe, érzelmbe és kognícióba
2023	**Számító birodalmak**, Kate Crawforddal	500 év technológia és hatalom	Az MI a kommunikáció, számítás, osztályozás és irányítás történelmi hálózatának legújabb szakasza
2025	**Fordított kontraszótár**, Vuk Ćosić és IOCOSE oldalával	generatív AI/LLM-ek	Maga a kulturális termelés a statisztikai nyelvi gépek alapanyagává válik
2026-tól kezdve	**Metabolic Machines**, Crawforddal	Az MI bolygó anyagcseréje	Megújult hangsúly a félvezetők, az energia, anyagok és ökológiai infrastruktúrák

Ezek a művek együtt alkotnak valami közel a **mesterséges intelligencia kritikus politikai gazdaságtanához és politikai ökológiájához**.

I. *Facebook Algoritmikus Gyár*: az elmélet kezdete

Joler érett MI-elmélete a SHARE Lab Facebook-kutatásaiból fakad, amely körülbelül 2016-ban kezdődött. *A Facebook Algorithmic Factory* rekonstruálta az **adatgyűjtés → algoritmikus feldolgozás → profilozás → célzás → monetizáció átmenetét**. Joler hangsúlyozta, hogy az algoritmusok közvetítik, hogy az információ láthatóvá válik, mely ötletek terjednek, mely bejegyzések tűnnek el, és így végső soron az egyén milyen társadalmi valóság verzióval találkozik.

A döntő koncepcionális lépés a hagyományos modell elhagyása volt:

A vállalat → szolgáltatás → fogyasztó.

Joler számára a platformkapitalizmus sokkal furcsább gazdasági kapcsolatot teremt. A látszólagos fogyasztó egyszerre termeli az árut.

A felhasználó tevékenységei – kattintás, lájkolás, megtekintés, közzététel, kommunikáció, hálózatokon való mozgás – olyan anyagokat generálnak, amelyekből a platform egyre részletesebb statisztikai profilokat épít fel. Ezek a profilok később reklámozásra, előrejelzésre, rangsorolásra és viselkedési beavatkozásra is használhatók. A SHARE Lab több száz algoritmust írt le, amelyek megpróbálják a felhasználókat egyre finomabb kategóriákba sorolni és a jövőbeli viselkedést leképezni.

Miért fontos ez későbbi MI-elmélete szempontjából

Itt három állandó jellemzője jelenik meg Joler gondolkodásának során.

Először is, a fekete doboz tézis. A társadalmi hatalom egyre inkább olyan rendszerekben rejlik, amelyek belső szabályai láthatatlanok azok számára, akik alávetettek nekik.

Másodszor, a gyári tézis. A digitális környezet nem csupán egy kommunikációs hálózat. Ez egy termelési eszköz.

Harmadszor, a fogyasztó inverziója. A fogyasztó magába a termelésbe épül.

Ez előrevetíti az *Anatomy of an AI System sokkal erősebb* megfogalmazását: az MI felhasználó egyszerre **fogyasztó, erőforrás, dolgozó és termék**.

Ez az egyik valóban gyümölcsöző ötlet, amely végigfut Joler egész életében.

II. *Egy MI rendszer anatómiája* (2018): Az MI nem jelentéktelen

A döntő áttörést Kate Crawford az *Antitomy of an AI System: The Amazon Echo című művében hozta az emberi munka, adatok és bolygók anatómiai* térképeként.

A látszólag egyszerű akció

emberi hang → Alexa → válasz

egy bolygó termelési rendszer látható csúcsaként rekonstruálják, amely bányákat, ásvány-feldolgozást, gyárakat, logisztikai hálózatokat, adatközpontokat, elektromos hálózatokat, kommunikációs hálózatokat, szoftvermérnököket, adatannotátorokat, raktári munkaerőt, fogyasztói képzési adatokat és végül elektronikus hulladékot foglal magában.

A munka **három alapvető kitermelő áramlást azonosít**:

1. **anyagi erőforrásokat;**
2. **emberi munkát;**
3. **adat.**

A legfontosabb újítás, hogy ezek nem különálló problémák. Egyetlen **integrált MI termelési rendszert** alkotnak.

1. MI mint geológiai technológia

Így az MI meglepő módon nem a matematikával, hanem a geológiával kezd.

A lítiumot, kobaltot, ónt, rézt és ritkaföldfém elemeket ki kell nyerni, finomítani, szállítani és átalakítani egy MI szolgáltatás létezése előtt. Az adatközpontok energiát fogyasztanak; a kommunikációs hálózatok kábeleket, routereket és műholdakat igényelnek; az eldobott berendezések elektronikus hulladékként térnek vissza a Földre.

Ezért a "felhő" látszólagos anyagtalansága ideológiai. Ami súlytalan számítási szolgáltatásnak tűnik, valójában rendkívül anyagi infrastruktúrákon alapul. Joler és Crawford szándékosan a térképüket a Föld kéregével kezdik és zárják.

Ez jelentős fogalmi változást eredményez:

Az MI-t nem szabad kizárólag szoftverként definiálni. Az MI egy ipari-anyagcsere rendszer.

Ez az állítás egyre fontosabbbá vált, ahogy az alapmodellek számítási igényei bővültek.

2. Az MI mint globális munkaerőhierarchia

A második réteg a bányászattól és gyári munkától kezdve a programozásig és az adat annotációig terjed.

Térképük egészen másképp jutalmazott munkaerő-formákat helyez el egyetlen globális termelési láncban: bányászok, összeszerelő munkások, logisztikai munkások, raktári alkalmazottak, kiszervezett adatcímkézők, mérnökök és vezetők. A látszólag autonóm AI tehát rengeteg nagyon emberi munkaerőtől függ.

Az elmélet ennek következtében egyértelműen marxista szerkezettel rendelkezik:

a természet → munka → árucikk → felhalmozódás.

De Joler fontos kiterjesztése, hogy **maguk az adatok és az emberi tapasztalatok is belépnek a kinyerés körébe**.

3. A felhasználó négyszeres státusza

Talán az Anatómiában *a legélesebb megfogalmazás* , hogy egy MI felhasználó lehet a következő dolgok:

- fogyasztó, aki megvásárolja az eszközt vagy szolgáltatást;
- egy **olyan erőforrás**, amely viselkedési és nyelvi adatokat szolgáltat;
- egy **munkavállaló**, aki visszajelzést ad, ami javítja a modellt;
- termék, amennyiben az eredmény anyagilag értékessé válik.

Joler számára tehát a hagyományos megkülönböztetés a gyártó és a fogyasztó között instabil lesz.

4. Az erőforrás-kitermeléstől az emberi kitermelésig

A záró érv magában foglalja azt a magot, amelyből *később az Új Extrektivizmus* kinőtt: a huszonegyedik századi kitermelés egyszerre nyúlik a bioszféra legtávolabbi részeibe, valamint az **emberi kognitív és affektív mélységekhez**. A gépi tanulási osztályozásokat ezután beépítik a lehetőségek, tudás és vagyon elosztásába.

Az MI kapitalizmusnak tehát két bányászati határa van:

lefelé a Földbe

és

befelé az emberbe.

Ez a kettős mozgás alapvető Joler számára.

III. *A Nooscope Manifested* (2020/21): Joler legmélyebb elmélete arról, hogy valójában mi is az MI

A Jolerhez köthető elméletileg legfejlettebb MI szöveg a *The Nooscope Manifested: AI as Instrument of Knowledge Extractivism*, amelyet Matteo Pasquinellivel írtak, és később az *AI & Society című folyóiratban jelentek* meg.

Kiindulópontja a mesterséges intelligencia kifejezésének támadása.

Joler és Pasquinelli azt állítják, hogy egy autonóm intelligens gépről beszélve arra ösztönöz, hogy képzeljük el egy kvázi-organizmust vagy "idegen elmét". A Nooscope ehelyett az **MI "szekularizálását"** javasolja: a gépi tanulást inkább úgy értik, mint eszközt, amely hatalmas mennyiségű adatból mintákat nagyít és kihúz.

A metafora a távcső.

Egy távcső nagyítja a távoli objektumokat.

Egy mikroszkóp apró tárgyakat nagyít.

A **Nooscope** hatalmas információs terekben nagyítja a mintákat.

Ez Joler és Pasquinelli egyik legeredetibb konceptuális konstrukciója.

1. Az alapvető gépi tanulási gyár

Gépi tanulást háromrészes összeszerelő sorra redukálnak:

Tanítási adatok → tanulási algoritmus → statisztikai modell.

A legfontosabb szociológiai állítás, hogy **az emberi társadalom minden szakaszban létezik**.

A képzési adatok korábbi emberi besorolásokat is megtestesítenek.

A címkék emberi taxonómiákat testesítenek meg.

A modell célok intézményi célokat képviselnek.

A hibák történelmi torzításokat reprodukálnak.

A kimeneteket később újra bevezetik a társadalomba.

Így az MI nem egy autonóm intelligencia, amely kívülről néz szembe a társadalommal. Ez statisztikailag visszahajlott társadalom.

IV. MI mint tömörítés: az egyik legerősebb tézis

A Nooscope alapvetően a gépi tanulást információtömörítésként értelmezi.

Egy hatalmas kép-, szó- vagy viselkedéskorpust tömörítünk statisztikai modellbe. A modell nem tartalmazza a világot. Matematikailag leegyszerűsített reprezentációt tartalmaz a kiválasztott adatokban felfedezett szabályosságokról.

Ez technikailag és politikailag is jelentős.

A tömörítés szükségszerűen magában foglalja:

A kiválasztás → egyszerűsítés→ → információvesztés közelítése.

Így a torzítások nem csupán véletlen hibák, amelyeket jobb mérnöki megoldással mindig meg lehet szüntetni. Némelyik az alapvető folyamatból ered, amely a heterogén valóságot statisztikai szabályszerűségké alakítja át.

A szerzők ennek következtében azt állítják, hogy a hatékony tömörítés érdekében végzett vállalati optimalizáció társadalmilag a **homogenizáció és a sokszínűség elvesztésének** motorjává válhat.

Ez közvetlenül egy erős szociálelméleti tézishez vezet.

V. A "múlt diktatúrája"

A gépi tanulási modellek statisztikai mintákat tanulnak korábbi adatokból.

Ezért a jövőt a múltból származó struktúrák értelmezi.

Joler és Pasquinelli felhívják a figyelmet az ebből következő **"új" felfedezetlenségére.**

Egy valóban példátlan esemény – szokatlan viselkedés, új metafora, váratlan úthelyzet, új kulturális forma – nincs statisztikai kategóriája a képzési eloszlásban.

Ezért a gépi tanulásnak intencionális tendenciája van arra, hogy az újdonságot korábban létező minták alapján értelmezze.

Legerősebb megfogalmazásuk az, hogy a gépi tanulás automatizálhatja egyfajta **"múlt diktatúráját"** a jelen felett.

Ennek óriási következményei vannak.

Adminisztráció során képes reprodukálni az örökölt osztályozásokat.

A rendőrségben a történelmileg rögzített rendőri mintákat a jövőbeli kockázatértékelésekké alakíthatja.

A foglalkoztatásban képes reprodukálni a történelmi toborzási struktúrákat.

A kultúrában képes stilisztikai szabályszerűségeket regenerálni.

A politikai kommunikációban statisztikailag sikeres narratívákat képes reprodukálni.

A társadalmi rangsorban a tegnapi egyenlőtlenségeket a holnap valószínűségeivé alakíthatja.

Az MI így körkör alakú szerkezetet hoz létre:

A múlt társadalom → adatokat → statisztikai modellt → döntéseket → jövőbeli társadalom számára.

A központi veszély tehát nem az, hogy az MI teljesen eltérne az emberiségtől.

Lehet, hogy túl **mélyen épít fel az emberiség felhalmozódott** múltjából.

VI. Osztályozás + előrejelzés = statisztikai kormányzás

Joler és Pasquinelli esetében a legtöbb gépi tanulási művelet két elemi függvényen keresztül érthető:

Osztályozás és **előrejelzés**.

A Classification a következőket kérdezi:

Milyen dolog/személy/esemény ez?

A Prediction a következő kérdéseket kérdezi:

A korábbi minták alapján mi várhatóan történhet ezután?

Amint ezek a műveletek a technikai rendszerekből társadalmi intézményekbe vándorolnak, a kormányzás mechanizmusaivá válnak. A Nooscope tehát egy kialakuló **osztályozási és előrejelzési** társadalomról beszél, ahol a gépi tanulás új megfigyelési és statisztikai adminisztrációs technológiává válik.

Itt kerül Joler a legközelebb Michel Foucault-hoz.

Foucault rekonstruálta, hogyan teremtettek börtönök, iskolák, kórházak, hadseregek és bürokráciák a normalizált egyéneket azáltal, hogy megfigyelik, rögzítik, vizsgálták és osztályozták őket.

A Joler-Pasquinelli érv a következő:

Az intézményi normalizáció egyre inkább számítási normalizációvá válik.

A regiszterből lesz az adatbázis.

A bürokratikus fájl az adatprofil lesz.

Az adminisztrátort részben a besoroló váltja fel.

A statisztikai populáció folyamatosan újraszámolt vektorokból és valószínűségekből álló populációvá válik.

VII. Az okságot korrelációval helyettesítve

Egy másik fontos episztemológiai figyelmeztetés az oksági összefüggést érinti.

A gépi tanulás rendkívül képes azonosítani a korrelációkat. De a korreláció nem teremti meg az oksági magyarázatot.

Ha az intézmények mégis statisztikailag feltérképezett szövetségeken keresztül kezdik irányítani az embereket, fennáll a veszélye, hogy:

"az X-hez hasonló emberek történelmileg Y-t csinálták"

csendben

"Ezt a személyt úgy kell kezelni, mintha Y-t akarná tenni."

A Nooscope ezért a prediktív rendőri munkát és az összehasonlítható alkalmazásokat a statisztikai következtetés politikai beavatkozássá alakításának tekinti.

Ez a racionalitás átalakulását jelzi:

az oksági magyarázat → valószínűségi előrejelzés.

És vele:

a végrehajtott cselekedet → beavatkozásért felelős felelősség a jósolt viselkedés alapján.

Ez potenciálisan mélyreható átalakulást jelenthet a jogban és a kormányzásban.

VIII. MI és munka: az automatizálás emberi munkaerőből épül

A Nooscope *munkaelmélete* sokkal tovább fejleszti az *Anatómia* meglátásait.

A szerzők elutasítják az MI képét, mint önmagát termelő gépet. A képzési adatok előkészítése, annotálása, értékelése, moderálása és javítása hatalmas számú emberi munkástól függ. Ezért az AI látszólagos autonómiája azon alapul, amit más kutatók **szellemmunkának** vagy **heteromációnak** neveznek.

De a mélyebb állításuk radikálisabb.

Egy tanítási adathalmazt tekinthető emberi **munka** diagramjaként.

Miért?

Mert amit az MI tanul, általában valamilyen emberi kompetenciából származik:

vizuális felismerést az őrök és a vezetők végeztek;

nyelvi osztályozást, amelyet írók és adminisztrátorok végeztek;

diagnosztikai értelmezést szakemberek végeztek;

fordítók végezték a fordítást;

Tudományos mintázatfelismerés, amelyet kutatók végeznek.

Az MI képzés tehát részben a következőkből áll:

egy emberi képesség megfigyelése, → termékeinek adatmá alakítása, → szabályosságok kinyerése, → a rendszeresség gépbe ágyazása.

A szerzők feltűnően foglalják össze a logikát: a gépi intelligencia a **munkamegosztásból** ered, míg egyik fő gazdasági célja a **munka automatizálása**.

A tömeghelyettesítés előrejelzése

A cikk kifejezetten előre jelezi a "hatalmas" hatást a foglalkoztatásra. Még provokatív példát is hoz, hogy ha egy mélytanulási rendszer képes egy rutinszerű vizuális felismerő feladatot körülbelül 99%-os pontossággal elvégezni, akkor a rutinmunka nagyon nagy része, amely erre a feladatra épül, potenciálisan automatizálhatóvá válhat, jogi és munkajogi korlátok alá téve.

Ezt a konkrét megfogalmazást óvatosan kell kezelni. **A feladatpontosság nem egyenértékű a szakmai pótolhatósággal.** A munkakörök feladatok, intézményi felelősségek, kivételkezelés, interperszonális funkciók és jogi elszámoltathatóság csomagjai vannak. A példát inkább úgy értik, mint strukturális érvük illusztrálása, mint megbízható kvantitatív munkaerőpiaci előrejelzésként.

A mélyebb előrejelzés azonban jelentős:

Az MI korábban megtestesített emberi kompetenciákat alakít át reprodukálható számítási tőkévé.

Ez valószínűleg sokkal jelentősebb lesz, mint bármely egyetlen elfoglaltság helyettesítése.

IX. *Új kizsákmányolás*: a megfigyelő kapitalizmustól a teljes kitermelés elméletéig

Az Új Extrektivizmus: Fogalmak és allegóriák összeállítása ismét kibővíti a rendszert.

Joler most azt kérdezi, mi történik, amikor gyakorlatilag **minden valóságdimenzió potenciálisan mérhető** lesz.

A kinyerés tárgya a következők lehetnek:

természeti erőforrásokat;

emberi munkát;

figyelem;

társadalmi kapcsolatok;

mozgás;

nyelv;

biometrika;

érzelem;

preferenciák;

hiedelmek;

kogníció;

DNS;

magának a testnek.

A térkép a digitális kapitalizmust úgy írja le, mint egy olyan infrastruktúrát, amely a **valóság aspektusainak kinyerésére, gyűjtésére, számszerűsítésére, elemzésére és összekapcsolására, valamint tőkévé alakítására** szolgál, koncentrálva a vagyont és a hatalmat néhány globális vállalatba.

1. "A test és az elme mint terület"

Ez Joler egyik legradikálisabb kiterjesztése a klasszikus politikai közgazdaságtannak.

A hagyományos gyarmatosítás földrajzi területeket foglalt el.

Az ipari kapitalizmus ásványokat, mezőgazdasági termékeket és munkaerőt termelt ki.

A digitális kapitalizmus új területeket fedez **fel az emberi életben**.

Az érzelem mérhető lesz.

A figyelem mérhető lesz.

A barátság grafikonadatmá válik.

A mozgás helymeghatározási adatmá válik.

A preferenciák valószínűséggé válnak.

Az identitás vektorokká válik.

Joler tehát hatékonyan értelmezi az adatfigurációt egy új bezárási formának.

Ami korábban egyszerűen élt, az a következik:

mérték→ elfogták, → bezárták, → átalakították eszközzővé.

X. "Affinitás a végtelenséghez"

Egy másik fontos fogalma az *Új Extrektivizmában*, amelyet Lyotardból kölcsönöztek, de Joler rendszerén átdolgoztak, a **végtelenséghez való affinitás**.

A digitális vállalatok állandó ösztönzést tapasztalnak arra, hogy bővítsék a felvételi lehetőségek skáláját és felbontását.

Több terület.

Több felhasználó.

Több érzékelő.

További adat.

Több paraméter.

Több viselkedési változó.

Több számítási kapacitás.

Több előrejelzés.

Az új határvonal tehát potenciálisan végtelen, mert szinte minden jelenség elméletileg adatokként írható le. Joler kifejezetten az új horizontot adatkitermelésként, gépi tanulásként és az információ újraszervezésének nevezi az ember-gép rendszerek kombinált rendszereivel.

Ez egy fontos dinamikai elméletet eredményez:

Az MI kapitalizmus strukturálisan terjeszkedő.

Ez nem csupán a meglévő szolgáltatások javítását jelenti.

Folyamatosan új jelenségeket keres, hogy kiszámíthatóvá tegye azokat.

XI. Egy különösen fontos Joler-sorozat: terület → erőforrás → monopólium → hiperhatalom

Egy 2024-es interjúban, amelyben *a Calculating Empires* című művéről beszélt, Joler szokatlanul világos sorrendbe tömörítette történelmi értelmezését.

Egy technológia új területet nyit meg.

A terület gazdaságilag kihasználhatóvá válik.

A színészek versengnek az irányításért.

Egy győztes vagy kis oligopólium domináns tulajdonjogot szerzett.

A dominancia a **hiperhatalom és kontroll** formájává válik.

Ezt visszatekintve alkalmazza az internetre: egy eredetileg decentralizált architektúra végül néhány óriási vállalat uralma alatt állt.

Ez valószínűleg Joler általános társadalomelméletének egyik legjobb megfogalmazása:

technológiai nyitás→ területteremtés→ árucikké válás→ verseny→ koncentráció→ monopólium/oligopólium→ politikai hatalom.

Az MI-re alkalmazva a következmény rendkívül jelentős.

A szűkös stratégiai területek a következők:

Képző korpra;

csúcskategóriás chipek;

felhőalapú számítástechnika;

adatközpontok;

alapozási modellek;

terjesztési platformok;

felhasználói interakciós adatok;

tudományos tudás;

Talán végül egész kognitív infrastruktúrát is felidéznek.

Így Joler jövőbeli félelme nem főleg:

Egy gép mindenhatóvá válik.

Ez a következő:

az MI stacket irányító szervezetek rendkívül erőteljessé válnak.

Ez a megkülönböztetés megkülönbözteti őt a klasszikus technológiai szingularitáselmélettől.

XII. *Birodalmak számítása* (2023): az érett történelmi szintézis

Kate Crawforddal Joler drámaian megnövelte az időskálát.

A Calculating Empires: A Genealology of Technology and Power Since 1500 öt évszázadot négy szervező kategóriában vizsgál:

Kommunikáció

Számítás

Osztályozás

Irányítás

A hatalmas térkép tartalmazza a kommunikációs technológiákat, interfészeket, algoritmusokat, modelleket, emberi számítógépeket, hardvereket, oktatást, érzelmi osztályozást, testeket, biometrikát, orvostudományt, rendőrséget, börtönöket, határokat, bürokráciát, gyarmatosítást, termelést, energiát, geológiát, légkört, építészetet, megfigyelést és katonai rendszereket.

Ez lényegében Joler kísérlete a **számítási teljesítmény makrotörténetére**.

1. A négy-C sorozat

Joler meglehetősen elegánsan magyarázza el a szerkezetet:

a kommunikáció médiát és adatokat szolgáltat;

a besorolás szervezi őket;

a számítás modellekké tömöríti őket;

A kontroll ezeket a modelleket visszavezeti a társadalmi valóságba.

Sématikus szempontból:

A világ → reprezentációja → osztályozás → számítás → modell → beavatkozás a világban.

Ezt nevezhetnénk Joler általánosított kibernetikai ciklusának.

XIII. A technológia soha nem semleges – de a Joler sem egyszerű technológiai determinista

Joler ismételten ragaszkodik hozzá, hogy a technológia nem semleges.

Egy interfész határozza meg, mely műveletek lehetségesek.

Egy osztályozási rendszer határozza meg, mely különbségek számítanak.

Egy algoritmus határozza meg, mely változók számítanak.

Az adatbázis határozza meg, mit lehet tárolni.

Egy modell határozza meg, mi válik kiszámíthatóvá.

De van egy fontos feltétel az elméletében.

A technológia maga egy szélesebb **gazdasági szabályok, tulajdonosi struktúrák, intézmények és ideológia rendszerébe** épül.

Példaként az internetet hozza fel. Technikai protokollja viszonylag decentralizált kommunikációt engedélyezett, mégis a kortárs kapitalizmus keretein belül az irányítás erősen koncentrálódott.

Így az ő álláspontja nem egyszerű:

A technológia dominanciát okoz.

Ez közelebb áll a következőkhöz:

A technológiai lehetőségek lépnek kölcsönhatásba a politikai-gazdasági struktúrákkal, és koncentrált tulajdonjog körül ismételten a felhalmozódás és ellenőrzés eszközeivé válnak.

Ez sokkal erősebb elmélet.

XIV. Az MI rejtett genealógiájának besorolása

A Calculating Empires *egyik legérdekesebb része* a besorolás genealógiája.

Joler és Crawford kapcsolatai a következők között találhatók:

természetes osztályozás;

faji taxonómiák;

antropometria;

biometrika;

IQ-teszt;

eugenika;

bürokratikus kategorizálás;

rendőrség;

határellenőrzés;

kortárs automatizált osztályozás.

Joler alapvető következtetése az, hogy a besorolás különösen veszélyessé válik, ha emberre alkalmazzák, mert a kategóriák könnyen vonnak intézményi következményeket.

Ebben az értelemben az MI példátlan **automatizálást jelent a besorolásban**.

Ahelyett, hogy egy bürokrata több ezer embert kategorizálna kategóriákba, a besorolók folyamatosan végrehajthatják a műveletet hatalmas népességeken keresztül.

Ezért a radikális jövőbeli lehetőség nem csupán a jobb mintázatfelismerés.

Ez a **társadalmi létezés folyamatos gépi osztályozása.**

XV. Az intézménytől az "adatiparig"

A Calculating Empires kifejezetten megfigyeli, hogy az intézmények az iskoláktól a börtönökig adatiparokká válnak, integrálva az átfogó elfoglalást és elemzést. Azt is állítja, hogy a tőke számára korábban elérhetetlen terek – az érzelmi kifejezéstől a külső űrig – már számítási kivonás tárgyává válnak.

Ha ez a pálya folytatódik, Joler logikája szerint a valószínű intézményi átalakulás a következő lenne:

iskolai → oktatási adatrendszer;

kórházi → biomedikai előrejelző rendszer;

foglalkoztatás → munkaerő-optimalizálási rendszer;

a rendőrség → előrejelző osztályozási rendszer;

a határ → biometrikus döntési rendszer;

média → figyelem-előrejelzési rendszer;

fogyasztói piaci → viselkedési előrejelző rendszer;

Adminisztráció → algoritmusan segített osztályozási rendszer.

Így az MI kevésbé válik különálló szektorrá, inkább egy **vízszintes infrastruktúrává, amely minden szektort áthatol**.

Ezért érdemli meg Joler gondolkodását átfogó képnek nevezni.

XVI. Generatív MI és a megosztott valóság válsága

A Birodalmak Számítása éppen akkor fejeződött be, amikor a generatív MI belépett a tömegkultúrába.

Joler megjegyezte, hogy emberek milliói gyorsan megváltoztatták a keresési, írási és képgenerálás módját. Három azonnali következményt emelt ki:

a hatalom koncentrációját;

hallucinációk és félretájékoztatás nagy léptékben;

egy közismert valóság eróziója.

Ez fontos új réteget ad a régebbi kivonási elméletéhez.

Előbb:

Az emberek adatokat állítanak elő, → platformok kinyerik azokat.

Generatív MI-vel:

Az emberi kultúra → képzési adatokat → modellezni → szintetikus kultúrát → emberi kultúrát.

Egy rekurzív hurok alakul ki.

XVII. *Reverse Contradictionary* (2025): a kulturális kivonás belép a generatív szakaszába

A 2025-ös *Reverse Contradictionary*, amelyet Vuk Ćosić és IOCOSE készített, művészibb és polemikusabb, mint *a Nooscope*, de elméletileg feltáró.

Kiindulópontja, hogy az emberiség kulturális termelésének hatalmas mennyisége – könyvek, cikkek, blogok, nyelv és beszélgetések – válik oktatóanyaggá az LLM-ek számára. A projekt támadja azt, amit a szerzők az emberi kulturális termelés statisztikai nyelvi generáló forrásokká

alakítanak, és szatirikusítja a technológiai "aranyozott kor" feltételezett kezdetét körülvevő ideológiát.

Az LLM-ekről szóló retorikát "forró tavakról és égető erdőkről" szóló retorikát szándékosan provokatív művészi nyelvezetként kell értelmezni, nem pedig kvantitatív környezeti értékelésként.

De fogalmi szempontból a kiterjesztés fontos.

Joler extrektivista sorozata a következőképpen változott:

Az ásványok → munka → adatok, → viselkedés → kogníció → kultúra.

XVIII. A feltörekvő *Metabolic Machines* projekt

Joler legutóbbi együttműködése Crawforddal azt sugallja, hogy elméletének következő szakasza még tovább halad az AI **anyagcseréjének nevezett** irányába.

2026-ban a Van Abbemuseum bejelentette, hogy közösen megbízza Crawford és Joler első filmjét, *a Metabolic* Machines-t, valamint Eindhoven félvezető-, optika- és lézeriparát, valamint a technológiai termelést megalapozó környezeti és társadalmi rendszereket vizsgálja. Egy kapcsolódó múzeumi program kifejezetten úgy írja le a munkát, mint az MI technológiák "anyagcsere" dimenzióját vizsgálja.

Ez teljesen összhangban van az *Anatomy of an AI System fejlesztésével*.

A valószínű fogalmi sorozat a következő lesz:

energia + víz + ásványok + chipek + munka + adat + számítás → mesterséges intelligencia képesség.

Más szóval, az intelligenciának van anyagcseréje.

Minél tovább bővül a számítás, annál inkább meg kell érteni az állítólag "virtuális" jelenséget az anyag és energia fizikai áramlásain keresztül.

XIX. Milyen radikális MI fejlesztést lát valójában előre Joler?

Itt szükséges egy fontos megkülönböztetés.

Elsősorban nem jósolja meg:

- tudatos gépek;
- autonóm gépi civilizáció;
- technológiai szingularitás;
- az emberi kihalás rekurzívan önfejlesztő AGI révén;
- Pontosan dátumozott átmenet a szuperintelligenciára.

Valóban, a *Nooscope* kifejezetten bírálja azt az elképzelést, hogy az MI-t független idegen kognícióként kellene elképzelni.

Radikális átalakulást vár a méret és az intézményi áthatolás révén.

Későbbi nyilatkozatai szerint az MI technológiai változásokat képes létrehozni, amely társadalmi méretben összehasonlítható a korábbi ipari forradalommal.

A teljes Joler-rendszerből legalább **nyolc jövőbeli transzformáció** rekonstruálható.

1. Az MI univerzális közvetítővé válik

A munka, tanulás, keresés, kommunikáció, kultúra létrehozása és döntéshozatal egyre nagyobb részeit modellek közvetítik.

A döntő átalakulás tehát a következő:

Az MI mint eszköz → MI mint infrastruktúra.

---

2. A kognitív munka válik az új automatizálási határvonalá

Az ipari kapitalizmus gépesítette a fizikai munkát.

Az MI kapitalizmus egyre inkább formálissá teszi, rögzíti és automatizálja a **szimbolikus és kognitív munka egyes** részeit.

Az írás, fordítás, keresés, osztályozás, kódolás, elemzés, tervezés és adminisztráció részleges automatizálásra válik sebezhetővé.

De Joler ragaszkodna hozzá, hogy az automatizálás alatt új rétegek maradjanak a látszólagos emberi munkaerő.

Így a valószínű jövő nem tiszta automatizálás, hanem a következő következő folyamatok mozgó keveréke:

automatizálás + emberi felügyelet + rejtett mikromunka + a dolgozók gépi kezelése.

3. A tőkekoncentráció gyorsul

Mert a határvidéki MI óriási koncentrációkat igényel:

főváros,

számítás,

chips,

energia,

adatok,

kutatási szakértelem,

felhőinfrastruktúra,

Joler terület-monopólium-modellje további koncentrációt jelent.

Pontosan ezért írja *le a Calculating Empires* az MI ipari átalakulását úgy, hogy kevesebb kézbe koncentrálja a hatalmat.

Valószínű, huszonegyedik századi politikai problémája tehát közel áll az **infrastrukturális oligarchiához**.

A cégek nem csupán árulnának termékeket.

Egyre nélkülözhetetlenebb rétegeket birtokolnának a kogníció és kommunikáció területén.

---

4. A statisztikai irányítás bővül

A besorolás és előrejelzés a következőkbe vándorol:

foglalkoztatás,

pénzügy,

biztosítás,

egészség,

oktatás,

rendőrség,

a vándorlás,

katonai ügyek,

közigazgatás.

Az emberek ezért egyre inkább géppel olvasható profilokon **keresztül találkoznak intézményekkel**.

A társadalmilag határozott "személy" fokozatosan a következőkre válhat:

Életrajzi Ember

és

statisztikai adatok dupla.

Az intézményi döntéseket egyre inkább az utóbbiakhoz irányítják.

---

5. A jóslat elkezdi formálni, amit előre jelez

Amint a jóslatok befolyásolják a hitelt, foglalkoztatást, rendőrséget, láthatóságot vagy oktatási lehetőségeket, már nem csupán a társadalmat figyelik.

Részt vesznek a gyártásában.

Így következik:

A jóslás → beavatkozás → megváltoztatott viselkedés → új adatok→ erősebb modell.

Ez potenciálisan önmegerősítő visszacsatolási hurkokat hoz létre.

Joler "múlt diktatúrája" itt különösen veszélyessé válik: történelmileg egyenlőtlen eloszlások gépi elvárásokká válhatnak, amelyek segítenek ezeket az eloszlásokat reprodukálni.

---

6. A kulturális homogenizáció lehetségessé válik

A Nooscope tömörítési érve, valamint az LLM kritikája a *Reverse Contradictionary-ben* új kulturális veszélyt sugall.

Ha a szintetikus szöveg, képek és zene egyre inkább a korábbi kulturális termelés statisztikai szabályszerűségeiből ered, és maga a szintetikus termelés később bekerül a jövőbeli képző korpusba, akkor az így következhet:

Kultúra → modell → szintetikus tenyésztés → jövőbeli képzési adatok → modell.

Joler elméleti aggodalma az lenne, hogy ez a visszacsatolási rendszer statisztikailag felismerhető formákat választ ki, és fokozatosan marginalizálja a valóban anomáliaikat.

A végső kulturális kockázat tehát nem feltétlenül a gépi eredetiség.

Ez **rekurzív normalizáció**.

---

7. A megosztott valóság kevésbé stabil lesz

A generatív rendszerek képesek meggyőző képeket, szövegeket, hangokat és videókat gyártani elhanyagolható marginális költséggel.

Joler kifejezetten összekapcsolja ezt a képességet félretájékoztatással és a megosztott valóság gyengülésével.

Egyénileg egyénileg javasolt rendszerekkel kombinálva a jövőbeli információs környezet így egyre inkább milliárdnyi személyre szabott előrejelzésből állhat, nem pedig korlátozott számú közös köznarratívából.

Ez potenciálisan mélyreható átalakulást jelent a közéleten.

---

8. Az AI ökológiai lábnyoma a képességeivel nő

Joler elmélete határozottan elutasítja azt a feltételezést, hogy az intelligencia növekedése egyre kevésbé lényegessé válik.

Az ellenkezője is előfordulhat.

Nagyobb modellképesség megkövetelheti a következőket:

több számítást;

nagyobb adatközpontokat;

több félvezető-gyártás;

nagyobb villamosenergia-fogyasztás;

víz hűtéshez;

több hálózati infrastruktúra;

nagyobb ásványi ellátási láncok.

A feltörekvő *Metabolic Machines* projekt megerősíti, hogy ez az anyag-ökológiai dimenzió továbbra is központi eleme Joler jelenlegi kutatásaiban.

XX. Joler legmélyebb jóslata: a küzdelem azon múlik, ki a kognitív infrastruktúra tulajdonosa

Szerintem ez a legjobb módja annak, hogy az egész elméletét szintetizáljuk.

A Bostrom-stílusú MI-elmélet esetében a döntő kérdés:

Mennyire lesz intelligens a gép?

Joler számára a fontosabb kérdés:

Ki birtokolja és irányítja azt a bolygóinfrastruktúrát, amelyen keresztül gépi intelligencia előállítható és alkalmazható?

Ez egy teljesen más jövőelméletet eredményez.

A központi veszély nem feltétlenül az, hogy egy mesterséges intelligencia kiszabadulna az emberi irányítás elől.

Lehet, hogy **az MI nagyon szilárdan emberi irányítás alatt marad – de rendkívül kevés ember és szervezet irányítása alatt áll.**

A fenyegetés tehát nem csupán a technológiai autonómia.

Ez **aszimmetrikus emberi erő, amelyet a technológia gyakorol.**

XXI. Joler elméleti szintézise

Az egész rendszere a következő láncra tömöríthető:

Természet

↓nyersanyagok és energia

Munkás

↓bányászat, gyártás, logisztika, annotáció, programozás

Emberi élet

↓Viselkedés, érzelem, kogníció, kommunikáció

Adatficálás

↓Felvétel és besorolás

Gépi tanulás

↓statisztikai tömörítés

Modellek

↓Osztályozás és előrejelzés

Intézmények

↓Foglalkoztatás, piacok, oktatás, rendőrség, határok, média, közigazgatás

Irányítás

↓rangsorolás, ajánlás, láthatóság, kizárás, beavatkozás

Felhalmozódás

↓Gazdagság és infrastruktúra hatalom

További terjeszkedés

↓új erőforrások, új adatok, új domainek

Ez egy visszacsatolási hurkot hoz létre:

Kivonás → számítás → előrejelzés → kontroll → felhalmozódás → kiterjesztett kivonás.

Ez, sokkal több, mint bármely konkrét diagram, **Joler általános MI társadalomelmélete**.

XXII. Mennyire eredeti ez az elméleti rendszer?

A Joler mindenképpen az elődekre épít.

A marxista örökség látható a kizsákmányolásban, a munkában, a többletben és a koncentrációban.

Foucault megjelenik a osztályozásban, normalizálásban, fegyelemben és intézményi tudásban.

Deleuzi eszmék jelen vannak a osztható/adatosított témában.

A megfigyelés-kapitalizmus elmélete átfedi a viselkedési adatok kezelését.

A média régészeti formája a történelmi genealógiát.

Posztkoloniális és dekolonalista elméletek határozzák meg az adatok és tudáskitermelés leírását.

A médiageológia hidat képez a számítás és a bolygóanyagok között.

Joler eredeti hozzájárulása azonban kevésbé abban rejlik, hogy minden komponens feltalálta, hanem inkább **olyan területeken összekötve, amelyeket a speciális elméletek általában elválasztanak**.

Sok megfigyelési elméleti szakember elemzi az adatokat, de figyelmen kívül hagyja a kobaltot.

Sok munkaelméleti szakértő elemzi a munkavállalókat, de figyelmen kívül hagyja a képzési adathalmazokat.

Sok számítástechnikus elemzi a modelleket, de figyelmen kívül hagyja a birodalmi történelmet.

Sok környezetelmélet elemzi a bányászatot, de figyelmen kívül hagyja az algoritmikus osztályozást.

Sok MI etika szakértője elemezi az elfogultságot, de figyelmen kívül hagyja a hajózást, az energiahálózatokat és a félvezető ellátási láncokat.

Joler egy térképen helyezi el őket.

Ezért az általa preferált intellektuális forma – a kritikai térképészet – nem csupán művészi módszer. Ez maga is elméleti módszere.

XXIII. Három különösen eredeti Joler-tézis

Ha a legjellegzetesebb hozzájárulásokat kellene kiválasztanom a munkáiból, három kiemelném.

1. Az MI egy bolygó anyagi rendszer, nem egy immateriális algoritmus

A gépi tanulási modell csak egy kis része annak a hatalmas stacknek, amely a bányáktól a globális munkaerőn át az adatközpontokig és az emberi tudáson át terjed.

2. A digitális kapitalizmus mindkét irányt egyszerre gyarmatosítja

Lefelé a geológiai Földbe, majd befelé az emberi kognícióba és affektív hatásba jut.

Ezért jön létre az új extrektivizmusság **egységes fogalma**.

3. Az MI legmélyebb politikai átalakulása a besorolás és előrejelzés privatizációja

A modern államok történelmileg bürokrácia, statisztika és intézmények révén osztályozták a lakosságot.

Egyre inkább olyan vállalatok fejlesztik és birtokolják a számítási besorolást, amelyek infrastruktúrája túllép a nemzeti határokon.

Ez történelmileg szokatlan áthelyezést jelent a részben kormányzati képességek számára, amelyek a **lakosság látása, kategorizálása és előrejelzése** voltak, magántechnológiai infrastruktúrákba.

Ez lehet Joler elméletének legjelentősebb politikai következménye.

XXIV. Korlátok és kritikák

Joler keretrendszere intellektuálisan erőteljes, de fontos gyengeségei is vannak.

Először is, a kitermelés fogalma túl átfogóvá válhat.

Ha a kobalt bányászatát, a böngészési viselkedés gyűjtését, nyelvi modellek képzését, reklámozást, biometrikus azonosítást és kulturális termelést mind kitermelési formáknak neveznek, akkor a fogalom mindent megmagyaráz, és így túl keveset különböztet.

Az *Új Extrektivizm* ereje – hatalmas integratív hatása – egyidejűleg elméleti veszélye is.

Másodszor, az elemzés normatívan aszimmetrikus.

Joler szándékosan a rejtett költségekre, a kizsákmányolásra, az átláthatatlanságra és a dominanciára koncentrál. Ennek következtében munkái sokkal kevesebb szisztematikus elemzést tartalmaznak az MI által létrehozott termelékenységről, orvosi előnyökről, tudományos felfedezésekről, hozzáférhetőségi technológiákról vagy decentralizált/nyílt forráskódú lehetőségekről.

Ezért kritikus **anatómia**, nem pedig kiegyensúlyozott jóléti elemzés.

Harmadszor, futurizmusa strukturális, nem előrejelző.

Néhány ilyen előbb:

numerikus forgatókönyvek,

dátumok,

valószínűségi becslések,

szektoronként történő pótlási előrejelzések,

A technológiai diffúzió formális modelljei.

Ezért kevésbé hasznos, mint a gazdasági előrejelzési tanulmányok annak becslésére, hogy hány munkahely tűnik el 2030-ig, de inkább arra a kérdésre**, hogy milyen társadalmi struktúrát építenek, ha az MI mindenütt elterjedt** lesz.

Negyedszer, *a Nooscope* megelőzte az alapmodell-forradalmat.

A gépi tanulás statisztikai tömörítésként való felfogása továbbra is megvilágosító, de a kortárs multimodális modellek, eszközhasználó rendszerek és egyre inkább ügynökös architektúrák olyan kérdéseket vetnek fel, amelyeket a 2020-as szöveg nem tudott teljes mértékben megválaszolni.

A legutóbbi *Reverse Contradictionary* és *Metabolic Machines* projektek azt mutatják, hogy Joler kiterjeszti a keretrendszert a generatív MI felé, de a teljes poszt-LLM elméleti átfogalmazás még nem jelent meg a The Nooscope szisztematikus formájában.

XXV. Általános értékelés

Az általad használt konkrét kritériumhoz – közép**- és kelet-európai szerzők, akik átfogó elméletet dolgoztak ki az MI társadalomra gyakorolt** következményeiről – Jolert meglehetősen magasra sorolnám, egy feltétellel.

Nem épített fel olyan klasszikus szociológiai rendszert, amely formában összehasonlítható lenne Parsons-sal, Luhmann-nal vagy egy szisztematikus társadalomfilozófussal. Elmélete ehelyett **diagrammai, genealógiai és politikai-gazdasági rendszer**. Az egyes művek egymás után rétegeznek egy intellektuális konstrukciónak:

A Facebook Algorithmic Factory fedezte fel az adatgyárat;

Mesterséges intelligencia rendszer anatómiája felfedezte bolygó anyagi testét;

A Nooscope rekonstruálta az episztemológiai gépezetét;

Az új extrektivizizmus általánosította gazdasági logikáját;

A Calculating Empires szolgáltatta az öt évszázados történelmi genealógiáját;

A Reverse Contradictionary kiterjesztette a kritikát a generatív nyelvre és kultúrára;

és **a Metabolic Machines** most tovább tolja az elméletet az egyre erősebb MI infrastruktúrák anyagi és ökológiai anyagcseréje felé.

Az eredmény egy rendkívül koherens állítás a közelgő MI társadalomról:

Az MI radikális jövője nem feltétlenül egy hirtelen tudatos szupergépként érkezik. Ez fokozatosan alakulhat ki, ahogy maga a társadalom is olyan infrastruktúrák köré szerveződik át, amelyek folyamatosan kihúznak, osztályoznak, előrejelzést nyújtanak és beavatkoznak – miközben az infrastruktúrák feletti irányítás egyre inkább koncentráltálódik.

Ez élesen megkülönbözteti Jolert a szingularitáselméleti elméleti szakértőktől. Központi jövőbeli forgatókönyve nem **a gépi szuverenitás az emberiség felett, hanem a** számítási kapacitás, gazdasági koncentráció és intézményi kontroll példátlan fúziójának lehetősége magában az emberi **társadalomban.**

Ebben a tekintetben elmélete talán közvetlenebb szociológiai, mint az AGI-ről szóló hagyományos vita: a döntő tényező nem csupán az, hogy az MI mennyire intelligenssé válik, hanem **az, hogy az infrastruktúrái mennyire teljesen befonódnak a munka, tudás, kultúra, közigazgatás, természet és hatalom** szervezésébe.

****

VI. Románia

****

Mihai Nadin elmélete

Kérjük, részletesen írja le és elemezze Mihai Nadin tanulmányait, amelyek során a mesterséges intelligencia társadalomra gyakorolt átfogó hatásait elemezte, beleértve az AI további radikális fejlődését, amelyet a következő években előrejelzések szerint láttak.

Mihai Nadin megérdemli, hogy az egyik legeredetibb kelet-közép-európai születésű gondolkodóként kezeljék, akik **valóban átfogó elméleti keretet fejlesztettek ki a mesterséges intelligencia megértéséhez**, bár megközelítése jelentősen eltér a ma már ismert technológiai szingularitás, szuperintelligencia vagy mesterséges intelligencia által vezérelt poszthumanizmus elméleteitől. Romániában született és tanult, később Németországban és az Egyesült Államokban is aktív volt, Nadin ötvözi a számítástechnikát, elektronikát, szemioitikat, filozófiát, esztétikát, ember-számítógép interakciót és az **előrejelző rendszerek** elméletét. A Texasi Egyetem Dallasban munkáját úgy jellemezte: átfogó elme, előrelátás, dinamikus rendszerek, mindenütt jelenlévő számítástechnika, vizualizáció és ember-technológiai kapcsolatok; megalapította az Előzetes Rendszerek Kutatási Intézetét.

Központi tézise meglehetősen élesen megfogalmazható:

Az MI forradalmi társadalmi következményei óriási lehetnek anélkül, hogy a gépek ugyanúgy intelligenssé válnának, mint az élőlények.

Ez választja el Nadint Kurzweiltől, Bostromtól és a legtöbb erős AGI forgatókönyvtől. Egyre látványosabb gépi teljesítményt és a számítás egyre mélyebb behatolását vár a társadalmi életbe, de azt állítja, hogy **a teljesítményt, a számítást, a jóslást és utánzást nem szabad összekeverni a jelentéssel, megértéssel, előrejelzéssel, kreativitással vagy élő intelligenciával**. 2019-es *Machine Intelligence: A Chimera című cikke különösen* erősen hangsúlyozza ezt a megkülönböztetést: szerinte a számítás zárt a jelentéstől; Az MI automatizálja az intelligenciával kapcsolatos tevékenységeket anélkül, hogy így intelligenssé válna.

Ez a tézis sokkal érdekesebbé válik, ha Nadin évtizedeken át kidolgozott nagyobb elméleti rendszerbe helyezik.

1. Nadin MI-elméletének fő alkotásai

Nadin egyetlen mesterséges intelligencia szociológiai monográfiáját sem írt, amely egy hagyományos "MI és társadalom" értekezéshez hasonlítana. Ehelyett egy integrált elmélet születik a digitális civilizáció, a szemiotika, az előrelátás, a számítás, az orvostudomány, a kreativitás, a fenntarthatóság és az MI témákban végzett egymást követő munkákból.

Munkássága	**Fő relevancia az MI és a társadalom szempontjából**
Az írástudatlanság civilizációja (1997)	Nagyszabású elmélet az ipari/műveltségi civilizációból egy magasan specializált, digitálisan közvetített civilizáció felé való átmenetről
Várakozás: A vég a cél (2003)	Az előrelátó cselekvés általános elmélete az élő rendszerek jellemzőjeként
"Várakozás és a mesterséges: esztétikák, etika és szintetikus élet" (2010)	Különbség a mesterséges számítás és az élő előrejelzés között; következmények az etikára és kreativitásra
"Az előrejelző profil" (2012)	Próbálkozások a várakozás formalizálására mint folyamatot
"Előrejelzés és számítás: lehetséges-e az előrejelző számítástechnika?" (2016)	Közvetlen vizsgálat arról, hogy a számítás képes túllépni a precident és előrelátóvá válhat-e
"Gépi intelligencia: Egy kiméra" (2019)	A gépi intelligencia hagyományos fogalmainak legszisztematikusabb kritikája
"Az MI célzása mozgó célra: egészség (vagy betegség)" (2020)	Az elmélet orvostudományra való alkalmazása
"Intelligencia bármilyen áron? Az MI meghatározásának kritériuma" (2023)	Az energia- és adathatékonyságot javasolták a gépi intelligencia értékelésének kritériumaként
"Nyers erőszámítás és intelligencia" (2023)	Az LLM-méretű számítás és erőforrás-igényes mesterséges intelligencia kritikája
Zavarja meg a tudományt: A jövő számít (2023)	Szélesebb körű kritika a redukcionista tudományról és az MI-ről; fenntarthatóság és előrejelző tudomány
2024–2025-ös interjúk és esszék	Deepfake-ek, kultúra, oktatás, hitelesség, energiafogyasztás, szabályozás és az AI jövőbeli társadalmi szerepe

A Civilizáció az írástudatlanság különösen fontos, mert megelőzte a kortárs generatív mesterséges intelligencia működését, mégis megadja azt a társadalmi elméletet, amelybe Nadin később az MI-t helyezi. A könyv 1997-ben jelent meg, és 881 oldalas terjedelmes. 2003-as *Várakozása* ekkor adja a későbbi elméletének egyértelműbb filozófiai alapját: a tervezés, a kreativitás és a tudáskeresés előrejelző jelenségek, amelyekben lehetséges jövőbeli állapotok befolyásolják a jelenlegi tevékenységet.

---

2. Az alapvető különbség: reakció, előrejelzés és előrejelzés

Minden más Nadin MI-elméletében egy alapvető különbségből fakad.

Hagyományos fizikai rendszerek esetén a sorozat lényegében a következő:

múlt → jelen → jövő.

Egy gép bemeneteket fogad, valamilyen struktúra szerint dolgozza fel azokat, és kimeneteket generál. A gépi tanulás sokkal összetettebbé, statisztikaibbá, adaptívabbá és átláthatatlanabbá teszi a struktúrát, de Nadin szerint alapvetően a múltbeli vagy jelenlegi adatokban gyökerezik.

Az élő szervezetek másképp működnek. Nemcsak a történtek miatt cselekszenek, hanem a **lehetséges jövőbeli állapotok miatt is**. Az éhség miatt az állat keres, mielőtt éhennem halna; a veszélyt elkerülik, mielőtt sérülés történik; az élőlények alkalmazkodnak, mielőtt a környezeti nyomás hatását befejezné; az emberek terveznek, képzelnek, alkotnak, ígérnek, félnek, reményt teremtenek és terveznek.

Nadin ezt **várakozásnak** nevezi.

Ezt nem szabad összekeverni az előrejelzéssel. A jóslás egy jövőbeli állapotot következtetik a múltbeli megfigyelésekből is. Az előrejelzés, ahogy Nadin értelmezi, azt jelenti, hogy a lehetséges jövők részt vesznek a jelen cselekvés meghatározásában. Enciklopédiájában, amely a prediktív és előrelátó számításokat vizsgálja, ezért megkülönbözteti a találgatást, várakozást, előrejelzést, előrejelzést és valóban előrejelző folyamatokat.

Ez a különbség azért fontos, mert a legtöbb modern MI, amely "előrelátásosnak" tűnik, valójában rendkívül fejlett előrejelzés.

Egy nyelvi modell valószínű folytatásokat jósol.

Az önvezető jármű előrejelzi a pályákat.

Egy orvosi mesterséges intelligencia előrejelzi a betegség valószínűségét.

Egy pénzügyi modell előrejelzi a piaci mozgásokat.

Egy ajánlóalgoritmus előrejelzi a preferenciákat.

Nedin kérdése sokkal radikálisabb: **Van-e a rendszer maga olyan jövő, amihez bármi számít?**

Egy élő szervezet számára a kudarc sérülést, éhséget, szaporodási képtelenséget vagy halált jelenthet. Az élőlény saját létezési állapotán belül cselekszik. Nadin azt állítja, hogy egy hagyományos gépnek nincs egyenértékű belső része. Céljai tervezőktől, operátoroktól, intézményektől vagy felhasználóktól származnak.

Ezért az előrejelzés nem csupán egy másik algoritmikus funkció az ő elméletében. Összefügg az **önfenntartással, az alkalmazkodással, a céllal és a** jelentéssel.

---

3. Az intelligencia inkább a jelentésről szól, nem pusztán sikeres teljesítményről

Nadin elméletének második pillére a szemiotika.

2019-es *Machine Intelligence: Egy Chimera című* művét kifejezetten kifejezi első szakdolgozatát:

"Az intelligencia a jelentésről szól, nem a mérésről."

Vegyük például a sakkot. Egy számítógép minden emberi sakkjátékost legyőzhet. Mégis, Nadin azt állítja, ez nem bizonyítja, hogy az ember úgy vesz részt a sakkban, mint egy ember.

Egy ember számára a sakk magában foglalhatja a versenyt, hírnevet, esztétikai elismerést, pszichológiai nyomást, kulturális memóriát, élvezetet, stratégiát, csalódást, kockázatot és személyes célt. A program olyan struktúrákat keres vagy tanul, amelyek lehetővé teszik a sikeres lépéseket. Ennek következtében az **egyenértékű vagy jobb eredmény nem jelenti azt, hogy az adott eredményt előállító folyamatok egyenértékűek**.

Az érv azonnal általánosítható a generatív MI-re.

Egy program szerkezetileg ekvivalens olyat tudhat létrehozni:

egy vers,

egy festmény,

egy bírói vélemény,

egy diagnózis,

egy szerelmes levél,

filozófiai érvet,

vagy egy zenei kompozíciót.

Nadin számára azonban a kérdés nem csupán az, hogy a megfigyelő meg tudja-e különböztetni az eredményt az emberi terméktől. Fel kell tenni a kérdést, mit jelentenek a **jelek a gyártó entitás** számára.

Ez teszi Nadint szinte a Turing-teszt hagyomány fordítójává. A Turing-megközelítés szándékosan a megfigyelhető teljesítményre koncentrál. Nadin azt állítja, hogy ez kockáztatja az utánzást a magyarázat helyett.

Innen jön az ismétlődő megkülönböztetése:

a mesterséges intelligencia mint teljesítmény ≠ az intelligencia mint élő folyamat.

A mélytanulás nem érvényteleníti a megkülönböztetést pusztán azért, mert teljesítménye rendkívül lenyűgöző. Az ő értelmezésében a statisztikai adatfeldolgozás hatalmas erejét mutatja be.

---

4. Miért nem változtatja meg a kortárs generatív MI Nadin diagnózisát

A nagy nyelvi modellek megjelenése eleinte elavulttá tehette korábbi érveit. Valójában Nadin az LLM-eket erősítőként értelmezte.

Az intelligencia bármilyen áron? című művében áttekinti a szimbolikus MI-től a kapcsolatosságig való átalakulást, valamint az új reményt, hogy a szimbolikus és neurális rendszerek ötvözésére irányul. Mégis, azt állítja, hogy az "AGI" mozgó célpont marad, amíg maga az intelligencia nem volt megfelelően meghatározva. Ahelyett, hogy pusztán a kimenet alapján értékelné az intelligenciát, azt javasolja, hogy vizsgáljuk meg **az eredmény elérésének eszközeinek megfelelőségét** – különösen azt a mennyiséget és energiát.

Ez szokatlan módon értelmezi a ChatGPT-szerű rendszereket.

Meglepő nyelvi kompetenciájuk azt mutatja, hogy hatalmas nyelvi szabályossági terek elsajátíthatók. De Nadin számára a modellméret növelése, az adatok tanítása és számítása nem feltétlenül mozgatja a rendszert a nem intelligencia és intelligencia közötti folyamatosan. Ehelyett egyre **meggyőzőbb mesterséges teljesítményt eredményezhet egyre erősebb számítások révén**.

Ez a megkülönböztetés Nadin egyik legfontosabb hozzájárulása a kortárs MI-elmélethez. Előre jelezte azt a problémát, hogy ugyanaz a megfigyelhető eredmény két nagyon eltérő értelmezést kaphat:

Értelmezés A: "A gép intelligenciát szerzett."

B értelmezés: "Felfedeztük, hogy ez a konkrét teljesítmény kevesebb mindent igényel, amit az emberek intelligenciának neveznek, mint ahogy korábban gondoltuk."

Nadin általában a második értelmezést választja.

---

5. Legeredetibb koncepciója: előrelátó és előrejelző MI

2016-os *tanulmánya, az Anticipation and Computation: Is Anticipatory Computing?* közvetlenül felteszi a kérdést.

Nadin elfogadja, hogy a gépek egyre korlátozottabb előadásokban képesek felülmúlni az embereket. A legfontosabb kérdése, hogy a gép képes valami szerkezetileg mást csinálni: nem csupán megnyerni egy meglévő játékot, hanem valóban új játékot indítson; nem csupán manipulálni egy bevált nyelvet, hanem részt venni az értelmek és új célok kialakításában.

A Springer-kötet, amelyben a tanulmány megjelent, kifejezetten a biológia, orvostudomány, kognitív tudomány, politika, társadalomtudomány, építészet és számítás területén helyezi el az előrejelzést.

Ez egy potenciálisan nagyon fontos taxonómiához vezet.

A prediktív MI szabályszerűségeket fedez fel a rendelkezésre álló vagy generált adatokban.

Az adaptív MI a visszajelzések szerint változtatja meg viselkedését.

A generatív MI új kombinációkat szintetizál a tanult reprezentációs terekben.

Az előrejelző intelligencia, Nadin erősebb értelemében, részt vesz saját jövőbeli lehetőségeinek és célainak meghatározásában.

Az első három egyértelműen létezik. A negyedik továbbra is megoldatlan probléma.

Így Nadin elmélete szigorúbb módot ad az AGI kérdés megfogalmazására, mint egyszerűen azt kérdezni, mikor fognak a számítógépek az embereknél felülmúlni a benchmarkokon.

---

6. MI és a munka átalakulása

Nadin társadalomelmélete korábban kezdődik, mint a jelenlegi MI-fellendülés. *Az írástudatlanság civilizációja* egy olyan gazdasági civilizációt ír le, amelyet egyre inkább a hatékonyság követelményei irányítanak: több termelés, gyorsabb forgalom, nagyobb specializáció és egyre kifinomultabb közvetítés technológiai ábrázolásokon keresztül.

Az MI pontosan ezt a tendenciát radikalizálja.

Az ipari gép a mechanikai erőt helyettesítette az izommunka helyett. A digitális számítás kezdetben a formális feldolgozást helyettesítette a rutinszerű szellemi műveletek helyett. A kortárs MI kiterjeszti a helyettesítést olyan tevékenységekre, amelyeket korábban kifejezetten kognitívnak tartottak:

fordítás,

osztályozás,

képértelmezés,

írás,

programozás,

adminisztráció,

tervezés,

diagnózis,

tervezés,

és döntéstámogatás.

Nadin jellegzetes figyelmeztetése nem elsősorban az, hogy a gépek autonóm személyekké válnak, és "elveszik a munkáinkat". Az a lényeg, hogy **a gazdasági szereplők egyre inkább átadják az emberi funkciókat gépekre, mert ezt a hatékonyság logikája jutalmazza**.

Ez mély különbség.

A kiszorítást előállító mechanizmus nem igényel tudatos robotokat. Csak olyan gépre van szükség, amelynek teljesítménye kellően olcsó, gyors, megbízható és skálázható.

Ennek következtében egy társadalom hatalmas MI forradalmon megy keresztül, még akkor is, ha Nadinnak teljesen igaza van abban, hogy maguk a gépek nem rendelkeznek valódi intelligenciával.

Ez talán elméletének legfontosabb szociológiai következtetése.

---

7. A mélyebb veszély: az emberek alkalmazkodnak a gépekhez

Nadin elmélete túlmutat a technológiai munkanélküliségen.

A technológiák nem csupán olyan funkciókat hajtanak végre, amelyeket korábban emberek végeztek. Megváltoztatják azokat az embereket, akik használják őket.

A számoló megváltoztatja az aritmetikai gyakorlatot.

A GPS megváltoztatja a navigációt.

A keresőmotorok megváltoztatják az emlékezetet.

A közösségi média megváltoztatja a kommunikációt.

Az ajánlórendszerek megváltoztatják a kulturális felfedezést.

A generatív MI megváltoztatja az írást, kutatást és problémamegoldást.

Így az MI visszacsatolási hurkot hoz létre:

Az emberek számítási eszközöket hoznak létre, → az emberek feladatokat delegálnak rájuk, → az intézmények átszerveződnek ezekre az eszközökre, → az emberek más képességeket szereznek, mert az átszervezett intézményekben élnek.

Ezért érthető Nadint leginkább nem hagyományos technológiai deterministaként, hanem **az ember-technológia együttátalakításának elméleti** alkotójaként.

A veszély tehát nem csupán a "gépi intelligencia növekedése". Ez az emberi **előrelátási kompetencia csökkenése is, mert egyre több olyan tevékenységet helyeznek ki, amelyek megítélést, orientációt és felkészülést igényelnek, a** rendszereknek.

Ez illeszkedik szélesebb diagnózisához, amely egy egyre inkább az azonnali megközelítésre irányuló civilizációról szól. 2007-es "Várakozás vonzóként" című művében a egyre dominánsabb "ÉN és MOST" irányultságot szembeállítja az "USA és JÖVŐ" szemlélettel, és összekapcsolja a technológiai-gazdasági hatékonyságot a csökkenő társadalmi koherenciával.

Az MI drámaian fokozhatná ezt a fejlődést.

---

8. Oktatás: a tárolt tudástól az ítéletig

Ez egy jellegzetes oktatási következtetést eredményez.

Ha a gépek kivételesen jóvá válnak az információ reprodukálásában, a szokásos kérdések megválaszolásában, szövegek összefoglalásában és jól meghatározott feladatok megoldásában, az oktatás, amely főként **az információszerzésre és reprodukcióra** szerveződött, egyre kevésbé védhetővé válik.

Nadin elmélete azt sugallja, hogy az oktatásnak pontosan azok felé kell haladnia, amelyeket ő a legkevésbé leegyszerűsíthetőnek tart számításra:

kontextuális megértés,

kérdések megfogalmazása,

ítélet,

értelmezés,

célkitűzés kialakítása,

kreatív újrakeretezés,

felelősség,

és izgatottság.

A különbség összefoglalható úgy, mint a **"Hogyan?" és "Miért?" közötti különbség.**

A gépek egyre erősebbé válnak abban, hogy "hogyan" adják a megoldásokat.

Nadin úgy gondolja, hogy az emberi szférának egyre inkább vissza kell nyernie a "miért".

Ez a téma egyértelműen megjelenik 2010-es mesterségesség, esztétika és etika elemzésében: a mesterséges termelésen túllépéshez valami élethez hasonló dolog felé kellene kezelni a kétértelműséggel és a "Miért?" dimenzióval, míg a számítási lehetőségek a jól definiált területen működnek.

Ez különösen relevánssá teszi Nadint a generatív MI oktatási vonatkozásai szempontjából. Ő szemszögéből nézve, ha a ChatGPT betiltása az iskolákból kihagyná a központi átalakulást. Maga a régi oktatási modell is működésképtelenné válik, amikor a gépek képesek sok hagyományos kimenetét reprodukálni.

---

9. MI, művészet és a közelgő hitelesség problémája

Nadin régóta dolgozik a számítás és a művészet metszéspontján, ami különösen érdekessé teszi a generatív MI-re adott reakcióját.

Nem tagadja, hogy a gépek esztétikailag lenyűgöző tárgyakat tudnak létrehozni. Ő inkább a **generáció és a teremtés közötti** különbséget alkotja.

A teremtés az ő értelemében magában foglalja az alkotó előrejelző állapotát: szándékokat, lehetőségeket, célokat, értékeket, kockázatokat és jelentéseket. Egy algoritmus képes megtanulni Beethoven, Rembrandt, Shakespeare vagy millió kortárs művészhez kapcsolódó struktúrákat, és formálisan hihető folytatásokat generálni. De ez nem jelenti azt, hogy a gép elfoglalja azt az egzisztenciális-kulturális kapcsolatot a munkájával, amely értelmessé teszi az emberi kreativitást.

Egy 2025-ös MI, kultúra és művészet témájú vitában Nadin ezért azt állította, hogy a társadalom kockáztatja az **utánzást kreativitásként**, és összekapcsolja az AI által generált tartalmat az autentiitás és fenntarthatóság problémáival.

Ez valami szélesebb jelentőséggel bír, mint egy filozófiai érv arról, hogy az MI "valóban teremt-e".

A szintetikus tartalom tömeggyártása átalakíthatja az **autentiitás társadalmi gazdaságát**.

Amikor képek, hangok, esszék és előadások elhanyagolható marginális költséggel készülnek, a szűkös elmozdulás a technikai gyártástól a következők felé tolódik:

eredet,

bízzunk,

Identitás,

szándékosság,

hírnév,

és igazolt emberi szerzőség.

Ez a következmény erősen összhangban van Nadin szemiotikus elméletével, még akkor is, ha nem pontosan ilyen kifejezésekkel fogalmazza meg.

---

10. MI, deepfake-ek és a közszféra átalakítása

Nadin a szintetikus médiát az egyik legnyilvánvalóbb közvetlen társadalmi veszélynek azonosítja.

Egy 2024-es interjúban hangsúlyozta, hogy az MI aggasztó része nem egy autonóm gép, amely kiszabadul az emberi irányítás elől, hanem az emberek, akik rendkívül erős eszközökkel rendelkeznek. Kitalált interjúkat és szintetikus képeket használt példaként: az MI óriási mértékben képes felerősíteni a hazugság, félrevezetés és befolyásolás képességét, mert az eredeti és mesterséges másolat közötti különbség nehezen megteremthető.

Ez fontos elmozdulás a szokásos "szuperintelligencia átveszi az irányítást" történettől.

Nedin modellje ehelyett azt javasolja:

Az MI radikális → a képviselet költségeinek csökkentését → a szintetikus reprezentációk radikális növekedését → a bizalom válságát → az ítélkezés és hitelesítés növekvő fontosságát.

A politikában ez azt jelenti, hogy az MI forradalmi hatásokat gyakorolhat, mielőtt bármi AGI-hoz hasonló létezne.

A kampánypropaganda, a személyre szabott meggyőzés, a szintetikus jelöltek nyilatkozatai, automatikus közösségi média kampányok és skálázható manipuláció nem igényelnek tudatosságot. Hatékony statisztikai modellekre van szükségük.

Nadin keretrendszere tehát egy fontos paradoxont azonosít:

minél kevésbé intelligens az MI, annál veszélyesebb lehet az emberi intelligencia és a hatalom túlzott tulajdonítása neki.

---

11. Etika: a gépek szabályokat hajthatnak végre, de nem generálnak erkölcsi jelentést

Etikaelmélete közvetlenül az előrelátásból fakad.

Az emberi erkölcsi cselekedet nem csupán a szabályok alkalmazása. Ez magában foglalja a cselekvés előtt előre elképzelt következményeket, a lehetséges jövőkért való felelősséget, a helyzetek értelmezését, célokat, értékeket és a tevékenység jelentését.

A 2024-es interjúban határozottan hangsúlyozza a lényeget: a gépeknek nincs belső erkölcsi értékük; egy művelet rendkívül bonyolult lehet anélkül, hogy megválaszolná az emberi kérdést, hogy miért kellene elvégezni.

Ez **nem** jelenti azt, hogy az MI rendszerek nem tudnának etikai korlátokat beépíteni.

Nyilvánvalóan megtehetik.

Egy rendszer programozható vagy betanítható úgy, hogy ne biztosítson bizonyos kimeneteket, optimalizálja a biztonsági kritériumok szerint, rangsorolja az alternatívákat normák szerint, vagy alkalmazza a konkrét jogi szabályokat.

Nadin erősebb pontja, hogy ez a normativás **származékos**. Valaki vagy intézmény választja ki a célokat, adatokat, korlátokat és szabványokat.

Ennek komoly következménye van az MI irányítása szempontjából: az erkölcsi felelősség nem "átruházható" egyszerűen a gépbe.

Minél autonómabbnak tűnik az MI rendszerek működése, annál fontosabbá válik meghatározni, mely emberi szereplők maradnak felelősek a céljaikért és következményeikért.

---

12. Az orvoslás mint döntő teszteset

A 2020-as *Aiming AI at a Moving Target: Health (or Disease) című könyve* különösen hatékonyan alkalmazza ezt a megkülönböztetést.

Nadin elfogadja a mélytanulás lenyűgöző eredményeit, de elutasítja mindkét szélsőséges álláspontot: "Az MI végül mindent megtehet az orvostudományban" és "az MI lényegében semmit sem tud hozzátenni."

A kérdése ehelyett így szól:

Mely orvosi komponenseket lehet megfelelően az MI-re bízni?

A cím "mozgó célpontja" kulcsfontosságú. Egy gép képes egy képet hatalmas adathalmazokban található minták szerint osztályozni. De az élő beteg nem csupán egy statikus tárgy, amely

osztályozásra vár. A beteg élettani változása; a kezelés megváltoztatja a beteget; viselkedés változik; környezet változik; a célok változnak; a biológiai rendszerek közötti kölcsönhatások változnak.

Az orvoslás tehát mindkettőt tartalmazza:

Nagyon kiszámítható előrejelző elemek,

valamint előrelátó, kontextuális, egyéni, élő folyamatokkal való kapcsolatot.

Nadin ezért az előrelátást az orvoslásba való beépítését követeli, nem pedig pusztán az adatfeldolgozás bővítését.

A kortárs jelentőség nyilvánvaló. Elmélete rendkívül erős MI-t jósolt a diagnosztika, mintafelismerés és döntéstámogatás terén anélkül, hogy arra a következtetésre jutna, hogy az orvos fogalmi szempontból redundánssá válik.

---

13. MI és tudomány: a "mérés megértés nélkül" veszélye

A Disrupt Science: The Future Matters *című művében* Nadin szélesíti kritikáját.

A modern tudomány rendkívüli eredményeket ért el redukció, mérés, determinizmus és bomlás révén. Ezek a módszerek rendkívül jól működnek élőlény nélküli fizikai rendszereken. Nadin azonban azt állítja, hogy az életre terjedés arra ösztönözte a tudósokat, hogy a egyre pontosabb méréseket összetévesztessék a egyre teljesebb megértéssel.

Az MI drámaian felgyorsítja ezt a fejlődést.

A gépi tanulás olyan összefüggéseket fedezhet fel, amelyeket egyetlen emberi kutató sem tudna kimutatni. Ennek következtében a tudományos gyakorlat egyre inkább a következőké válhat:

adatgyűjtés → gépi mintafelismerés → sikeres előrejelzés,

anélkül, hogy ekviáns növekedés lenne az oksági vagy értelmes megértésben.

Nadin nem utasítja el ezt a hatalmat. Ellenvetése az, hogy teljes tudással azonosítja azt.

A Disrupt Science ennek következtében azt a "második tudományforradalomnak" **nevezett cselekvést javasol** , amely az élő rendszerek előrelátó jellegére összpontosít. A kiadó előadása kifejezetten kimondja, hogy a digitális gép korszaka, különösen az MI, jelentős lehetőségeket teremt, de mély kockázatokat is teremt, és az adatok mérését és feldolgozását a jelentés megvalósításával állítja szembe.

---

14. Az energia-és adat kritériuma: Nadin szokatlan intelligencia-mérője

Az egyik legeredetib, legutóbbi javaslata a 2023-as *Intelligence at Any Price?* című tanulmányokban jelenik meg. és *a Brute Force Computation and Intelligence*.

Tegyük fel, hogy két rendszer ugyanazt a feladatot teljesíti.

Ehhez hatalmas képzési készletek, hatalmas számítástechnikai infrastruktúra és nagy energiafelhasználás szükséges.

Egy élő szervezet ugyanezt a funkciót éri el egy kis energia és kevés tapasztalat felhasználásával.

Nadin azt állítja, hogy a hatékonysági különbség nem véletlen. Ez bizonyíték az intelligencia természetére.

Ezért javasolja, hogy összehasonlítsuk a természetes és mesterséges teljesítményt azzal, hogy megkérdezi, mennyi **adat és energia** szükséges a cél eléréséhez.

Ez átalakítja az MI vitát.

Ahelyett, hogy:

Felülmúlhatja a számítógép az embert?

Nadin kérdezi:

Milyen információs és energikus áron teljesíti felül az embert?

Ennek fontos következményei vannak a jövőbeli MI számára.

A számítási kapacitás százszoros növekedése, amely tíz százalékos javulást eredményez, a mérnöki sikert jelentheti, miközben Nadin mércéjével mérve nem közelíti meg az **élő intelligenciát**.

Így a skálázás kétértelmű: képes az MI-t képességűbbé tenni, miközben az ő kritériuma szerint kevésbé intelligenciává válik.

---

15. A fenntarthatóság tehát az intelligencia definíciójának részévé válik

A legtöbb elméleti szakember számára az MI intelligencia és a környezeti fenntarthatóság külön témák.

Nadin számára ezek konceptuálisan kapcsolódnak.

Egy élő rendszer folyamatosan működik véges energetikai és anyagi korlátok alatt. Előrelátó szervezete lehetővé teszi, hogy túlélési célokat érjen el anélkül, hogy minden lehetséges alternatívát keresne.

A brute-force számítás másképp működik: a korlátozott megértést sokkal több adat feldolgozásával és sokkal több számítással kompenzálhatja.

Ez eredményezi azt, amit **Nadin anti-skálázási tézisének** nevezhetnénk:

A több számítás nem automatikusan több intelligenciát jelent.

A jövőbeli MI tehát két lehetséges fejlesztési irányba néz szembe.

Az egyik irány a egyre nagyobb modellek, adathalmazok, adatközpontok és számítási kiadások.

A másik minőségileg eltérő architektúra, amely sokkal több relevanciát tud kinyerni sokkal kevesebb információból.

Nadin a másodikat közelebb tartja a valódi intelligenciához.

Ez rendkívül előrelátó eleme elméletének, mert a figyelmet a benchmark pontszámokról a **számítás ökológiájára tereli**.

---

16. Nadin hiszi, hogy az AGI és a szuperintelligencia jön?

Itt a válasznak szokatlanul pontosnak kell lennie.

Nem—nem a szokásos kurzweiliai vagy bostromi értelemben.

Nadin nem jósolja meg azt a közeli pontot, amikor a hagyományos számítógépes MI hirtelen átlép egy mennyiségi küszöböt, és emberhez hasonló általános intelligenciává válik.

Valójában mély szkepticizmussal tekint az AGI szabványos fogalmára.

2019-es tanulmánya "gépi intelligenciát" kimérának nevezi, míg 2023-as cikke az AGI-t nem definiált vagy mozgó célpontnak tartja. Egy 2023-as nyilvános esszéjében tovább ment, és azt állította, hogy Turinghoz és Gödelhez kapcsolódó korlátok az általános gépi intelligenciát "kimérává" teszik a jövőbeli technológiától függetlenül.

Ezt az utolsó érvet azonban el kell választani Nadin elméletének erősebb részeitől.

Gödel befejezetlenségi tételeinek következtetése a gépi intelligencia lehetetlenségére rendkívül **vitatott**. A Stanford Filozófiai Enciklopédia megjegyzi, hogy a gödeli érvek, amelyek azt állítják, hogy az emberi elme szükségszerűen túllép a gépeken, komoly ellenvetésekkel

szembesülnek, és hogy széles körű konszenzus van abban, hogy az egyenes anti-mechanista következtetés meghiúsít.

Ezért két Nadin tézisről teszek különbséget:

A gyengébb tézis – miszerint a benchmark teljesítmény, statisztikai előrejelzés és szintaktikai előrejelzés önmagukban nem mutatják be az értelmet vagy az emberi-egyenértékű intelligenciát – filozófiailag erőteljes és továbbra is nagyon releváns.

Az erősebb tézis – miszerint a matematikai logika bizonyítja az AGI lehetetlenségét – sokkal nehezebben védhető, és nem szabad matematikailag bizonyítottnak tekinteni.

---

17. Nadin tényleges előrejelzése: Az MI továbbra is elterjedt lesz

Az AGI határozott elutasítása **nem** jelenti azt, hogy Nadin technológiai stagnációt jósolna.

Épp ellenkezőleg.

Már 2010-es tanulmányában megfigyelte, hogy a mesterséges intelligencia, a mesterséges evolúció, az empátia, a mesterséges kultúra és a mesterséges élet olyan mértékben bővül, hogy a gépek mélyen uralt jövője már nem pusztán sci-fi volt.

2024-ben az interjúztató azt kérdezte tőle, hogy az MI egyre jelentősebb részévé válik az emberi életben; Nadin válasza nem tagadta ezt az irányt. Elsősorban az aggodalom az volt, hogy az emberek és kivételes erővel bíró eszközökkel lépnek kapcsolatba, nem pedig a gépek önálló élő ügynökké válnának.

Ez azt jelenti, hogy jövőbeli forgatókönyve valójában **szociológiailag radikálisabb, mint ahogy a gépi intelligencia filozófiája elsőre sugallná**.

Előre látja azt a világot, ahol a gépek nem kell, hogy emberré váljanak, mert az emberek átszervezik társadalmukat gépek köré.

---

18. Milyen lenne a radikális MI fejlesztése Nadin elmélete alapján

Fontos itt megkülönböztetni **Nadin explicit állításait** elméleti rendszerének extrapolációjától a 2020-as évek végéig.

Elmélete szerint a legnagyobb átalakulás nagyjából 2026 és a 2030-as évek eleje között nem feltétlenül egy mesterséges elme születése lenne. Ez valószínűbb, hogy egyre képesebb, egyre alkalmasabb MI fokozatos beépítése lenne szinte minden társadalmi koordinációs rendszerbe.

Az adminisztratív munka gépi közvetítésűvé válik.

A tudományos kutatás egyre inkább MI-alapú segítséget igényel.

Az orvosi diagnózis egyre előrejelzőbbé válik.

Az oktatás személyre szabottvá és számításszerűvé válik.

A kulturális termelés szintetikus tárgyakkal árasztja el.

A szervezeti döntések algoritmusan előkészítetteké válnak.

A közlekedés fokozatosan autonómtá válik.

A megfigyelés mindenütt jelen van és következtethetővé válik.

A politikai kommunikáció egyre szintetikus és személyre szabott lesz.

Az ember-számítógép interakció kevésbé láthatóan "számítógépszerűvé": ahelyett, hogy szándékosan programokat üzemeltetnének, az emberek egyre inkább olyan környezetekben élnek, amelyeket folyamatosan értelmeznek és módosítanak a számítási rendszerek.

Ez a forgatókönyv sokkal természetesebben követi Nadint, mint egyetlen látványos "szingularitást".

Forradalma inkább **diffúzió, mint ébredés**.

---

19. A szakmák közelgő átalakulása

Ez lehetővé teszi egy fontos következtetést az elméletéből.

A legkiszolgáltatottabb munkahelyek nem feltétlenül azok, amelyek a legkevésbé intelligenciát igényelnek hagyományos pszichológiai értelemben. Azok azok, amelyek eredményei a legkönnyebben **formalizálhatók, értékelhetők és reprodukálhatók**.

Így a könyvelés, adminisztráció, fordítás, dokumentumkészítés, ügyfélszolgálat, rutinszerű jogi elemzés, programozás, radiológiai osztályozás, jelentés és kutatási segítségnyújtás jelentős része rendkívül automatizált lehet, anélkül, hogy a gépek olyasmit tudnának, amit Nadin élő intelligenciának nevezne.

Maguk a szakmák ezután változnak.

Egy ügyvéd átléphet az első vázlat elkészítésétől a releváns probléma meghatározásáig és az MI által generált vázlat érvényesítéséig.

Az orvos elmozdulhat a szokásos minták felismeréséről a változó beteg ellentmondásos információinak integrálására.

Egy tudós átléphet a szakirodalom kereséséről és az alternatívák számításáról a kérdések keretezésére és gépi által generált felfedezések értelmezésére.

A tanár átléphet az információ átadásáról a megértés értékelésére és az ítélőképesség fejlesztésére.

A paradoxon az, **hogy minél erősebbé válik az MI a "Hogy?" kérdésen, annál értékesebb lehet gazdaságilag valóban emberi kompetencia a "Miért?"** kérdésnél.

Ez valószínűleg a legeredményesebb módja annak, hogy Nadin elméletét a következő évekre is kiterjesszük.

---

20. Egy lehetséges emberi kompetenciaválság

De Nadin elmélete az ellenkező eredményt is sugallja.

Az emberek nem használhatják az MI-t pusztán azért, hogy megszabaduljanak a rutinfeladatoktól. Lehet, hogy abbahagyják a gépek által felváltott képességek fejlesztését.

A navigáció egyszerű precedenst kínál: a kitartó GPS használata csökkentheti a saját térbeli orientáció kialakításának szükségességét.

A generatív MI hasonló hatásokat okozhat a következőkben:

írás,

emlék,

érvelés,

számítás,

kutatás,

tervezés,

és végül a döntéshozatal.

Az eredmény szokatlan történelmi fordulat lett volna.

Az MI fokozatosan képes lesz, mert elszívja az emberi kulturális termékeket, míg az emberek egyre függőbbé, mert abbahagyják a képességek gyakorlását, amelyek ezekben a termékekben megtestesülnek.

Nadin elmélete ezért inkább a **függőségre irányítja a** figyelmet, nem csak a pótlásra.

A veszély nem az, hogy az MI emberré váljon.

Az a lényeg, hogy az emberek mesterséges rendszerek köré szervezik át életüket, míg korábban az emberi autonómiához tartozó funkciókat ritkán látják el.

---

21. Miért különösen fontos Nadin a generatív MI megértéséhez

A generatív MI három régóta fennálló megkülönböztetését rendkívül relevánssá teszi.

Először is, **szimuláció kontra azonosság**. Egy rendszer képes szimulálni az intelligens tevékenység termékeit anélkül, hogy feltétlenül megismételné az alapul szolgáló folyamatot.

Másodszor, **az előrejelzés kontra a várakozás**. A rendkívül pontos valószínűségi termelés nem hoz létre automatikusan olyan entitást, amelynek saját, jövőorientált egzisztenciális érdekei vannak.

Harmadszor, a **jelfeldolgozás kontra jelentés**. A jelek rendkívül koherens módjai manipulálásának képessége feltárja azt a filozófiai kérdést, hogy a jelek jelentenek-e valamit magának a rendszernek.

Ezért az LLM-ek sem egyértelműen, sem nem cáfolják Nadint.

Drámaian megerősítik, milyen messzire **lehet fejlődni a teljesítmény egységes intelligenciaelmélet** nélkül.

Hogy ezek cáfolják-e erősebb állítását, miszerint a jelentés és az előrelátás élő szervezettséget igényel, továbbra is nyitott filozófiai és tudományos kérdés.

---

22. Társadalomelméletének legproblémásabb része: szabályozás

Nadin politikai szabályozással kapcsolatos álláspontja sokkal vitathatóbb, mint az egyszerű MI fogalmakat érintő kritikája.

A 2024-es interjúban, amikor szintetikus félretájékoztatással szembesült, határozottan ellenezte a kormányzati szabályozást, és többször is az **oktatást** javasolta: a polgároknak tájékozott ítélőképességet kell fejleszteniük, nem pedig benyomások és propaganda irányítják őket.

Ez következetesen következik az előrejelző filozófiájából: a kompetens emberi ügynököknek inkább az ítélőképességet kell fejleszteniük, nem pedig egy másik reaktív bürokráciára bízniuk a felelősséget.

De ez egyben gyengeség is.

Az oktatás javíthatja az egyéni ellenálló képességet a megtévesztéssel szemben, ugyanakkor sok MI kockázat kollektív cselekvési problémákat foglal magában, amelyeket az egyéni felismerés nem tud megoldani: adattulajdonítás, monopóliumhatalom, átláthatatlan automatizált döntések, infrastruktúra-biztonság, felelősségvállalás, diszkriminatív rendszerek vagy hatalmas szintetikus információs műveletek.

Így elfogadható Nadin oktatáshoz való ragaszkodása, miközben elutasítjuk majdnem kettősségét az oktatás és a szabályozás között.

---

23. Ahol Nadin a legerősebb, hol a leggyengébb

Nadin elméleti rendszerének legerősebb része, hogy nem hajlandó azonosítani a **megfigyelhető gépi teljesítményt az azt előállító folyamat természetével**.

Ezt könnyebb volt figyelmen kívül hagyni a hagyományos programok korszakában. A kortárs generatív modellek esetében a probléma sokkal komolyabbá vált. Egy MI rendszer most olyan teljesítményeket tud létrehozni, amelyeket az emberek ösztönösen intelligenciának, kreativitásnak és megértésnek tulajdonítanak. Nadin évtizedeken át figyelmeztetett, hogy maga a tulajdonlás független igazolást igényel.

Második nagy erőssége az előrejelzés/előrejelzés megkülönböztetése. A kortárs MI túlnyomórészt statisztikai kapcsolatok elsajátítására és következtetések készítésére szerveződik. Ha azt kérdezik, hogy ez valódi várakozásnak számít-e, egy mély fogalmi problémát tár fel, amelyet az antropomorf nyelv elrejt.

Harmadik erőssége az intelligencia és **az erőforrás-elégesség közötti kapcsolat**. Ez arra kényszeríti az MI elméletet, hogy ne csupán azt mérlegelje, mit ér el egy rendszer, hanem hogy milyen pompásan éri el azt.

Negyedik erőssége a szociológiai: a technológiák megváltoztatják az őket használó embereket és intézményeket. Ezért nem lehet az MI-t pusztán közvetlen gépi képességekkel értékelni.

Leggyengébb állításai abszolút megfogalmazásai. Nem bizonyították, hogy csak a biológiai rendszerek lehetnek értelmesek, és hogy Gödel és Turing bizonyítják, hogy az AGI lehetetlen. Az adat- és energia kritériuma feltáró hatékonysági mérőszám, de valószínűleg nem elegendő az intelligencia teljes definíciójaként. Egy pazarló rendszer még mindig rendelkezhet bizonyos intelligens képességekkel, ahogy egy biológiailag hatástalan ember sem szűnik meg intelligensnek.

---

24. Nadin kontra a szingularitáselméleti elméleti szakemberek

A különbség végül sématikus módon fejezhető ki.

Kurzweil **számára** a számítási kapacitás növelése végül gépi intelligenciát eredményez, amely meghaladja az emberi intelligenciát.

Bostrom **esetében** a elég erős általános intelligencia szuperintelligenciává fejlődhet, ami irányítási/igazítási problémát eredményez.

Nadjon számára az alapfelvetés már eleve téves, ha a számítási teljesítményt egyszerűen az intelligenciával azonosítják. A számítás növelése rendkívül képes mesterséges rendszereket generálhat, miközben az algoritmikus feldolgozás és az élő előrejelző jelentés közötti alapvető különbséget érintetlenül hagyja.

A paradoxon az, hogy Nadin ezért **nem** jósolja meg kevésbé zavaró jövőt.

Más mechanizmussal jósolja meg a zavarokat.

Kurzweil radikális jövője a következő:

A gépek egyre inkább emberhez válnak.

Nadin radikális jövője közelebb áll a következőkhöz:

Az emberi társadalom egyre inkább gépek köré szerveződik.

Ez talán végső soron a legfontosabb különbség.

---

25. Nadin átfogó értékelése, mint átfogó MI társadalomelméleti szakember

A közép- és kelet-európai szerzők szélesebb csoportja kontextusában, akiknek MI-elméleteit vizsgáltuk, Mihai Nadint szokatlanul magasra helyezném elméleti **eredetiség és rendszerszintű teljesség** tekintetében.

Az eredetisége nem abban rejlik, hogy pontosan megjósolja, hány foglalkozás fog eltűnni, vagy mikor következik be egy technológiai mérföldkő. Szándékosan elutasítja ezt a technológiai előrejelzési stílust.

Ehelyett egy integrált koncepcionális architektúrát épít össze, amely összekapcsolja:

A szemiotika → → élő rendszereket jelent → előrejelzést, →intelligenciát → számítást, →ember-gép interakciót → társadalmi szervezettséget → etikát → kultúrát → orvostudományt → oktatást → fenntarthatóságot.

Ez jóval több, mint egy átlagos MI-kritika.

Legeredetibb állításai szerintem ezek a következők:

Az intelligenciát meg kell választani a nagy teljesítményű adatfeldolgozástól; az előrejelzést meg kell választani az előrejelzéstől; az élő intelligencia jelentés és önmaga által generált célok között működik; a technológiai tárgyak emberi felhasználóikat és feladataikat is megváltoztatják; az MI társadalmi forradalma nem igényel AGI-t; és figyelembe kell venni az előadás energetikai/információs költségét, amikor eldöntik, hogy megérdemli-e az intelligensnek nevezni. Ezeket a felvetéseket a munkája során fejtjük ki, nem pedig egyetlen elszigetelt publikációban.

A legmélyebb következmény a következő évekre ezért kissé ellentmondásos. **Nadin egyszerre lehet AGI-szkeptikus, és az rendkívül radikális MI által vezérelt társadalmi átalakulás elméletalkotója.** Nem várja el, hogy a döntő esemény egy számítógép "felébredése" legyen. A döntő esemény az észlelés, memória, előrejelzés, reprezentáció, kommunikáció, választás és végül az ítélkezés elemei fokozatos átruházása a számítási tárgyakra – és az ebből következő átalakulás során az emberek, intézmények és kultúrák egyre inkább rájuk támaszkodnak.

Ebben az értelemben a központi kérdése az MI korszakhoz végső soron nem az, **hogy "Mennyire lesznek intelligensek a gépek?"** Hanem az, **hogy "Mi történik az emberi előrelátó intelligenciával, amikor a civilizáció olyan gépek köré szervezi magát, amelyek egyre nagyobb arányban képesek utánozni az intelligens teljesítményt?"**

Ez nem egy hagyományos szingularitás előrejelzés, hanem a valóban radikális jövő, amely Nadin elméleti rendszerében található.

****

VII. Kelet-európai elméletalkotók eredeti tézisei és állításai

Kérjük, emelje ki a csatolt kötetből azokat az MI-elméleti téziseit és az egyes szerzők állításait, amelyeket eredeti újdonságoknak lehet nevezni a szerző társadalomra gyakorolt elméleti ismeretei szempontjából, amelyeket a világ minden táján kifejeztek.

A csatolt kötet szokatlanul megterhelő mércét állít fel: nem számít átfogó elméletnek a foglalkoztatás, etikák, jog, orvostudomány vagy oktatás területén végzett ágazati munkát, hacsak a szerző nem kínál koherens MI-fogalmat, társadalmi átalakulás oksági mechanizmusát, többszférát átfogó társadalmi elemzést és egy radikálisabb jövőbeli MI vagy ember-gép integráció felé való kiterjesztést. A kötet maga kifejezetten az MI társadalomra gyakorolt hatásainak átfogó elméleteinek van szentelve Kelet-Európában 2000 és 2026 között.

Egy módszertani minősítés elengedhetetlen. **A "globálisan eredeti" nem azt jelenti, hogy bizonyítani lehet, hogy soha nem jelent meg hasonló mondat a hatalmas nemzetközi irodalomban.** Ez külön, kimerítő bibliometrikus vizsgálatot igényelne sok nyelven keresztül. Én az erősebb, akadémiailag védhető jelentést használom: egy tézis valódi **elméleti újdonságra alkalmas**, ha egy jellegzetes mechanizmust, fogalmi architektúrát vagy meglepő következtetést vezet be, amelyet nem lehet egyszerűen Bostrom, Kurzweil, Zuboff, Harari, Castells, McLuhan, Stiegler, Foucault, transzhumanizmus, mesterséges intelligencia szerinti elmélet stb. ismertebb elméleteire redukálni. Ezen a szigorúbb teszten megkülönböztetném az **eredeti fogalmakat** a **meglévő fogalmak eredeti szintéziseitől**.

A kötet lényegében ugyanarra a következtetésre jut: négy konstrukció kerül a legmagasabb kategóriába—Pokol, Dukaj, Joler és Nadin—különösen szorosan követi őket Zybertowicz; Kelemen és Greguric–Vertovšek egy további eredeti csoportot alkotnak, míg Csepeli eredetisége inkább az átfogó szintézisben rejlik.

1. Béla Pokol — nagyon magas eredetiség

Pokol legerősebb állítása a nemzetközi elméleti eredetiségre nem az **a** jóslat, hogy az MI rendkívül erőssé válik. Ez nyilvánvalóan osztozik Kurzweillel, Bostrommal és egy nagy futurisztikus irodalommal. Az ő eredetisége abban rejlik, hogy megváltoztatja mind az **elemzés** egységét, mind végül a **probléma ontológiáját**.

Először is, a kollektív szuperintelligencia mint a társadalom tulajdonsága, nem pedig egy egyéni gépé. A nemzetközi irodalom már tartalmazta a kollektív intelligenciát, a globális agyat, az elosztott kogníciót és az ember–gép hálózati fogalmakat. Pokol jellegzetes lépése, hogy ezt egy **általános szuperintelligencia szociológiájának** alapjává teszi: milliók vagy milliárd ember, algoritmus, adatbázis, érzékelő és intézmény válhat annyira kognitívan integrálttá, hogy a szuperintelligens entitás gyakorlatilag a teljes társadalmi kognitív architektúra lesz. A kötet

kifejezetten azonosítja ezt az átmenetet az egyéni mesterséges elmétől a társadalom kognitív integrációjába, mint a négy legerősebb elméleti újítás egyikét.

Ez eredetib, mint egyszerűen azt mondani, hogy "az internet kollektív intelligenciát teremt." A erősebb állítás a következő:

Emberi intelligencia + intézményi intelligencia + gépi intelligencia + mindenütt jelen lévő adatintegráció → emergent társadalmi szuperintelligencia.

Ennek megfelelően egy társadalom szuperintelligenciává **válhat, mielőtt** bármi lenne egyetlen bostromi szuperintelligens ügynök.

Másodszor, a technológiai intelligencia és a társadalmi érték-értékű intelligencia közötti különbség. Pokol azt állítja, hogy a rendkívül gyorsított problémamegoldás és a technológiai intelligencia nem feltétlenül kell egyenértékű képességgel járni, hogy történelmileg kifejlődött igazságosság, hűség, méltóság, felelősség, szolidaritás és ellentmondó társadalmi értékek komplexumainak beilleszkedjenek. A jellegzetes veszély tehát nem feltétlenül egy gonosz gép, hanem egy **technológiailag szuperintelligens, mégis társadalmilag értékvak intelligencia**. A kötet ezt kiemeli Pokol saját, egyedi állításai közül.

Nyilvánvaló affinitások vannak itt a nemzetközi összehangolás problémájával. Ennek ellenére a specifikus szociológiai megfogalmazás jellegzetes: az értékek nem csupán egy véges preferenciák listája, amelyeket egy objektív funkcióba kell beilleszteni; történelmileg felhalmozódott és belsőleg ellentmondó struktúrák, amelyek társadalmi intézményekbe ágyazódnak.

Harmadszor, a hartmann-féle "evolúciós újraindítás" tézis valószínűleg Pokol legszokatlanabb állítása. Ahelyett, hogy egyszerűen az MI-t ötödik szintre adná a fizikai, biológiai, pszichikus és emberi-intellektuális valóság fölé, Pokol azt állítja, hogy az autonóm mesterséges intelligencia közvetlenül összekapcsolhatná az intellektuális szintet **a fizikai-anyagi alapanyaggal**, megkerülve a biológiai életet és a pszichikus tudatosságot. Az evolúciós szekvencia már nem csupán lenne

Anyag → élet → pszichés → emberi intelligencia

de elágazhat

Az anyag → mesterséges intelligencia → önirányított technológiai evolúció.

A kötet kifejezetten ezt nevezi az evolúciós újrakezdés rendkívül eredeti lehetőségének a fizikai alapoktól. Ez túlmutat Kurzweil szingularitásán vagy Bostrom intelligenciarobbanásán: az autonóm MI-t úgy értelmezi, mint egy **második földi intelligencia genealógiát indítani**. A kötet végső soron azt a lehetőséget, hogy az emberiség válik közvetítővé, amelyen keresztül az intelligencia már nem válik veleszületett biológiai szinten, Pokol legradikálisabb elméleti hozzájárulásának nevezi.

Negyedszer, az információbőség növelheti az intellektuális hierarchiát, mintsem megszüntetheti. Pokol megfordítja a központi digital-utópikus elvárást: az információhoz való univerzális hozzáférés nem eredményez univerzális kognitív egyenlőséget. Az MI aránytalanul megsokszorozza azoknak a szereplőknek a termelékenységét, akik már képesek absztrakcióra,

szintézisre és kifinomult kérdésekre; ennek következtében a tudás formális demokratizálása együtt létezhet egy **kognitív-technológiai szuperelit** megjelenésével.

Ötödször, a munka utáni piac–demokrácia inverziója nagyon eredeti, mint származtatott szociológiai tézis. A legtöbb nemzetközi technológiai munkanélküliségi elmélet megáll az alapon, az újraelosztásnál vagy az új munkahelyeknél. Pokol azt kérdezi, hogy a szinte teljes automatizálás megszünteti-e a *piaci koordináció funkcionális okának egy* részét: ha az emberi munka nem lesz döntő, szűkös termelő bemenet, a bérösztönzők és a munkaerőpiacok elveszíthetik korábbi központi szerepüket. Ezután az ellentmondásos második állítást vezeti le, miszerint **minél kevésbé válnak nélkülözhetetlenek a piacok, annál nélkülözhetetlenebbé válhat a demokrácia**, mert a kollektíven kezelt automatizált termelési kapacitás feletti politikai irányítás rendkívül jelentőssé válik.

Hatodik, a 2016-os "Kérdős" tézis és későbbi "tudásbányászat" kiterjesztése egy szokatlanul előrelátó elméletet alkot az MI által közvetített intellektuális munkáról. Amikor a válaszok számítási szempontból bőségessé válnak, a szűkös szellemi erőforrás problémák felépítésére, termékeny kérdések sorozatainak feltevésére, válaszok értékelésére és nézőpontok újrakombinálására irányul. A kötet a 2016-os tézistől a későbbi ChatGPT-alapú "tudásbányászat" módszer közötti folytonosságot azonosítja Pokol egyik jellegzetes állításaként.

Globális értékelés: a legerősebb állítások az igazán új elméletekre szerintem a **társadalmi kollektív szuperintelligencia**, a **hartmann-i evolúciós újraindítás** és a **piac-demokrácia inverziója szinte teljes automatizálás** alatt. Az értékintelligencia és a megkérdőjelező tézisek is jellegzetesek, bár több nemzetközi megfelelőjük is van.

---

2. Jacek Dukaj — nagyon magas eredetiség

Dukaj eredetisége abban rejlik, hogy a szokásos kérdést, "mikor lesznek a gépek olyan intelligensek, mint az emberek?" egy több ezer évre nyúló történelemmel helyettesítik: **az emberi civilizáció fokozatosan külsőleg képezte a Logost – szimbolikus feldolgozást, érvelést és végül döntéshozatalt – az egyéni embereken kívülre.**

A kötet hat különösen eredeti elemet azonosít elméletében.

A Logos külső irányába helyezett tézis a nyelvet, az írást, a formális intézményeket, a számítást és az MI-t egy történelmi pályára helyezi. Ennek elődei is voltak – McLuhan, Ong, kiterjesztett elme elméletek, Stiegler memória exteriorizációja és több technológiai filozófia –, de Dukaj jellegzetes hozzájárulása, hogy közvetlenül összekapcsolja ezt a történelmet az **emberi szubjektivitás jövőbeli redundanciájával**.

Legeredetibb mechanizmusa az **önkéntes deszubjektifikáció sikeres optimalizáláson** keresztül. Az embereket nem kell az MI meghódítani. Fokozatosan delegálják a döntéseket, mert az AI navigációja, diagnózis, párosítása, befektetése, adminisztrációja és előrejelzése jobban működik. Minden delegáció egyediben racionális, míg az összesített következmények forradalmiak:

jobb algoritmikus döntés → önkéntes delegáció → még jobb algoritmikus teljesítmény → az autonóm ítélkezés → csökkenő kulturális értéke .

A kötet pontosan hangsúlyozza azt a lehetőséget, hogy a jólétet maximalizáló MI hatékonyabban szüntetheti meg a szubjektivitást, mint a kényszerítés.

Ez hozza létre Dukaj különösen fontos **"láthatatlan szingularitás" tézisét.** Nem kell olyan napnak lennie, amikor "az MI átveszi a hatalmat". A navigáció, keresés, felvétel, befektetés, orvosi értelmezés, kutatás és adminisztráció egyenként kerül megbízásra. Végül a gyakorlati szuverenitás átvándorolt anélkül, hogy azonosítható alkotmányos hatalomátadás történt volna.

Egy másik, nagyon eredeti tézis a **rekurzív kulturális hurok**:

az emberi kultúra → MI képzés → mesterséges intelligencia által generált kultúra → technológiailag átalakított emberek → új képzési adatok → új MI kultúrá.

A kulcspont, hogy az emberi preferenciákat már nem lehet független, fix szabványként kezelni, ha a technológiai környezet maga hozta létre ezeket a preferenciákat. Ez jóval mélyebb, mint a szokásos szűrőbuborék tézis.

Végül Dukaj azt állítja, hogy **a fejlettebb MI kevésbé emberivé válhat, mint emberibbé**. Az emberhez hasonló nyelvi interfészek átmeneti alkalmazkodást jelenthetnek, miközben az emberek továbbra is ügyfelek és felügyelők maradnak. Amint a gépi kogníció túlmutat a főemlős agy, testesülés, munkamemória, érzelmek és halandóság korlátain, kognitív struktúrái egyre inkább eltérhetnek a miénktől.

Globális értékelés: Dukaj egyes komponenseinek fontos nyugati elődeivel rendelkeznek, de ezek integrációja az **önkéntes posztszubjektivitás és a láthatatlan szuverenitás átadásának** elméletébe rendkívül jellegzetes. Ez az egyik legerősebb jelölt a kötetben az igazán új globális MI társadalomelmélet számára.

---

3. Vladan Joler — nagyon magas eredetiség, de gyakran együttműködő

Jolernél a tulajdonlásnak pontosnak kell lennie: legerősebb elméleti eredményei közé tartoznak **a Joler–Kate Crawford** vagy **Joler–Matteo Pasquinelli** konstrukciók, nem pedig kizárólag egyéni tézisek.

Alapvető újítása, hogy elutasítsa az MI mint anyagi "gépi agy" ábrázolását. Az *AI rendszer anatómiájában* az MI egy hatalmas anyagláncgá válik, amely összeköti a bányászatot, ásványokat, globális munkaerőt, a gyártást, a logisztikát, az energiát, az adatközpontokat, programozókat, adatmunkásokat és felhasználókat. A *Nooscope* és *az új extrektivizmus* esetében ez egy tudáskitermelés politikai gazdaságtanába terjeszkedik.

A kötet három különösen eredeti Joler-tézist különít el.

Az első az MI **, mint bolygói anyagrendszer**. Ezért az MI geológiával, munkarendszerrel és egyre inkább ökológiai anyagcserével rendelkezik. Ezt ma már nemzetközileg is elismerik fontos hozzájárulásként a kritikus MI vizsgálatokhoz.

A második a **kettős extrektivizmus**: a digitális kapitalizmus egyszerre terjed **ki és lefelé a Földbe**, valamint **befelé az emberi kognícióba és affektívába**. Az ásványok, energia és víz az egyik póluson nyerik ki; a viselkedés, a nyelv, a figyelem, az érzelem és a tudás a másikon. Az emberi test és az elme maga is új, kitermelő "területekké" válik.

A harmadik a **besorolás és előrejelzés privatizációja**. A korábbi államok hatalmas hatalmat halmoztak fel népszámlálások, statisztikák, nyilvántartások és bürokratikus osztályozások révén. A fejlett MI példátlan képességeket ad magánkézben lévő infrastruktúráknak a népességek folyamatos osztályozására és előrejelzésére. Ez tehát a társadalom "önlátásának" egy részét áthelyezheti a közintézményekből a transznacionális magáninfrastruktúrák felé.

Két további Joler–Pasquinelli-tézis érdemel hangsúlyozást. A **"múlt diktatúrája"** azt jelenti, hogy a történelmi szabályszerűségekre kiképzésen alapuló rendszerek hajlamosak a jövőbeli társadalmi struktúrákból történő osztályozást építeni:

A múlt társadalom → adatokat → statisztikai modellt → intézményi döntéseket → jövőbeli társadalom számára.

Az MI így képes reprodukálni a történelmet, nemcsak azért, mert a programozók előítéletekkel rendelkeznek, hanem mert maga a jóslás strukturálisan a statisztikai visszatérésre irányul.

A **Nooscope** tézis pedig a gépi tanulást nem mesterséges elmeként, hanem statisztikai észlelés és tömörítés eszközeként kezeli: a kötet emlékezetes megfogalmazásában az MI **statisztikailag visszahajlott társadalommá válik**.

Globális értékelés: Joler munkája valószínűleg abban a kötetben is igaz, ahol a globális eredetiség állítása legkönnyebben alátámasztható, mert *maga az Anatomy of an AI* System és *a Nooscope* nemzetközileg befolyásos konceptuális alkotásokká vált. Fő újdonsága az ontológia változása: **az MI nem csupán szoftver, hanem egy bolygói társadalmi-anyagi infrastruktúra, amely kinyerés, osztályozás és hatalom áll.**

---

4. Mihai Nadin — nagyon magas eredetiség

Nadin szinte pontosan a hagyományos szingularitáselmélet koncepcionális fordítóját hozza létre. Alapvető állítása, hogy **az MI radikálisan képes átalakítani a civilizációt, még akkor is, ha a gépek soha nem válnak élő-emberi értelemben intelligenssé**.

Az elméleti alap a **jóslás és az előrejelzés közötti különbség**. A jóslat a múltbeli és jelenlegi állapotokról származó adatokból vezet le valószínű jövőbeli állapotokat. Nadin elmélete szerint az élőlények olyan lehetséges jövőkkel viszonyítva cselekszik, amelyek fontosak a létezésük folytatásához. A jelentés, a cél és a várakozás tehát nem egyenértékben összehangolható a egyre pontosabb statisztikai számításokkal.

Nadin legmélyebb társadalomelméleti inverziója ebből ered. A probléma nem az, hogy a számítógépek fokozatosan emberszerűvé válnak; lehet, hogy **az emberi intézmények, és végül maguk az emberek is gépszerűvé válnak**. A szervezetek a valóságot mérhető változókká alakítják le; az algoritmusok ezeken az egyszerűsített reprezentációkon működnek; az eljárásokat az algoritmusok köré alakítják át; az emberek alkalmazkodnak hozzájuk viselkedésüket; és így a valóság egyre könnyebbé válik a számítási rendszerek számára. A kötet kifejezetten ezt nevezi a regionális irodalom egyik legerősebb eredeti tézisének.

Így következik:

Összetett emberi valóság → a számítási egyszerűsítés → intézményi áttervezés → viselkedési alkalmazkodás → még géppel olvashatóbb valóság.

Ez rendkívül fontos alternatíva az AGI-központú elméletekkel szemben. A gép nem feltétlenül szerez jelentést; **a társadalom feláldozhatja a jelentést azért, hogy számítási szempontból irányítható** legyen.

Nadin egy **skálázásellenes tézist is** kidolgoz: a nagyobb számítás, adat és energiafelhasználás nem automatikusan jelent nagyobb intelligenciát. Az intelligenciát részben annak alapján kell megítélni, hogy egy ügynök gazdaságilag azonosítja, mi számít, nem pusztán a mérföldkőeredmény eléréséhez használt nyers számítási erő alapján.

A kötet végső értékelése legeredetibb állításait azonosítja: a nagy teljesítményű feldolgozás és intelligencia, az előrejelzés és előrejelzés, a külső célok és az életcél közötti különbségeket, valamint az állítást, hogy az MI megváltoztatja a felhasználókat és intézményeket, és hogy a társadalmi MI forradalom nem függ az AGI-től.

Globális értékelés: maga az előrejelzésnek is vannak előzményei, különösen az előrejelző rendszerek elméleteiben, és a számítási elme elme kritikáinak nyilvánvalóan nagy múltra tekinthetők be. Nadin különleges újdonsága abban rejlik, hogy ezeket a filozófiai különbségeket társadalmi **számítási adaptációs elméletté** alakítja: a gépeknek nem kell emberi jelentést szerezniük ahhoz, hogy az emberek átalakítsák a civilizációt olyan gépek köré, amelyekből ez hiányzik.

---

5. Andrzej Zybertowicz — magastól nagyon magas eredetiségig

Zybertowicz egyes komponensein láthatóbb nemzetközi elődei vannak, mint az előző négy szerzőé, de ő megalkotja a kötet egyik legeredetibb **oksági** rendszerét.

Első feltűnő tézise az **"új sötét korszak" inverziója**. A tudományos fejlődés hagyományosan növeli az emberi megértést. Az MI megfordíthatja ezt a történelmi kapcsolatot:

Az emberi tudomány → mesterséges intelligencia által generált tudás → emberfeletti összetettség → az emberi érthetőség romlása.

Így a civilizáció technikailag fejlettebbé válhat, miközben kognitívan sötétebbé válhat az emberi lakói számára. A politikai következmények mélyek: a formális emberi szuverenitás fennmaradhat, miközben a jelentős episztemikus szuverenitás eltűnik.

A második a **kétgörbe tézis**. A kritikus küszöb nem kizárólag attól függ, hogy az MI egyre képesebbé válik. Az egyik görbe a gépi képesség növekedését jelképezi; a másik pedig az emberi képességek potenciálisan csökkenését jelenti a tartós figyelem, mélytanulás és összetett ítélőképesség tekintetében. Az irányítás elvesztése előfordulhat, amikor a két görbe keresztezi – még akkor is, ha maga az AI nem lett istenszerű AGI. A kötet ezt kifejezetten Zybertowicz egyik legjellegzetesebb mechanizmusaként mutatja be.

A harmadik a **digitális előfeltételes tézis**, amelyet legerősebb, globális megkülönböztető hozzájárulásának tartom. A fejlett MI nem lép be abba a társadalomba, amely évtizedekkel korábban feltalálta a számítógépet. A közösségi média, a digitális platformok és az algoritmikus környezetek már átalakították a figyelmet, a hatalmi viszonyokat, a társadalmi kommunikációt és az intézményi kompetenciát. Ezért:

Az emberek digitális technológiát építenek, → a digitális technológia megváltoztatja az embereket, → átalakult emberek fejlett MI-t építenek, → az MI egy már átalakult társadalomra hat hatással.

A kötet maga ezt a rekurzív modellt azonosítja Zybertowicz legerősebb eredeti hozzájárulásaként.

A negyedik a **techno-csapda**: a vállalati és állami szintű racionalitás civilizációs szintű irracionalitást eredményez. Minden vállalat és állam magánéletben előnyben részesítheti az óvatosságot, de nem lassíthat biztonságosan, ha a versenytársak folytatódnak. Az MI veszélye tehát részben a versengés társadalmi struktúrájából fakad, nem magának az MI architektúrájából.

Ötödik értékes korrekciója az **összhangolásról** szól. Az MI "emberi értékekkel" való összehangolás azonnal alkotmányos kérdést vet fel: kinek az értékei, ki választott ki, milyen legitim eljárással? Ez átalakítja az összehangolást egy tisztán mérnöki kérdésből legitim kollektív normaválasztás elméletévé.

Globális értékelés: a technológiai verseny, az összehangolási politika, az átlátszóság és a technológiai függőség mind nemzetközi precedensekkel rendelkezik. A erősebb eredetiség abban rejlik, hogy beépülnek az MI rekurzív felvetéséhez, **amely egy olyan civilizációba lép, amelyet kognitívan és intézményileg átalakítottak – és potenciálisan gyengítettek – a megelőző technológiai** folyamatok által.

---

6. Jozef Kelemen — magas eredetiség, különösen a korai korai megjelenése miatt

A Kelemen legnagyobb jelentősége, hogy több, a generatív MI után rendkívül releváns állítás már a kilencvenes és a 2000-es évek elején is kidolgozott.

Legérdekesebb tézise a **kölcsönös antropológiai hurok**:

Az emberek az emberi intelligencia modelljei alapján építenek gépeket, → gépek megtestesítik ezeket a modelleket, → az emberek elkezdik magukat a gépekből származó kategóriák alapján értelmezni.

A gép tehát nem csupán technológiai termék; antropológiai tükörvé válik.

Második megkülönböztető újítása a **négy szakirányú technikaizációs** tézis, amelyet rendkívül korán fogalmazott meg: az intelligencia technikaizálása, a környezet technikizálása, a test kiborgizációja, valamint az érzelem és kreativitás technikaizálása. Utólag tekintve különösen figyelemre méltó a kreativitás és az érzelmi élet bevonása a mai generatív mesterséges intelligenciák előtt.

Harmadszor, Kelemen az **intelligenciát és a kreativitást heterogén ember–gép rendszerek emergendens tulajdonságaiként** kezeli. Az elemzés helyes egysége nem feltétlenül "a kreatív ember" és a "kreatív gép" lehet. A kreativitás a rekurzív kölcsönhatásukból alakulhat ki. Ez különösen hasznos elméleti keret a mai prompt–generáció–választás–felülvizsgálati ciklusokhoz.

Negyedszer azonosítja az adattól az információ → hosszú távú mozgását → **tudás** → **bölcsesség** között, és azt kérdezi, mi történik, amikor az első három egyre technikaibbá válik. A tudás/bölcsesség alapvető megkülönböztetés kétségtelenül nem világszerte új – a DIKW családnak és a régebbi filozófiának nyilvánvaló előzményei vannak –, de Kelemen alkalmazása egy fontos MI-paradoxont teremt: a tudás sikeres automatizálása társadalmilag értékesebbé teszi a nem automatizált kontextuális ítéletet. A kötet kifejezetten összefoglalja ezt a négy hozzájárulást.

Globális értékelés: Kelemen rendszere nem teljesen példátlan a kibernetika, Minsky, Brooks, poszthumanizmus és elosztott kogníció elmélet szempontjából. Eredetisége **korai integratív előrejelzés** egy olyan világra, ahol egyre lehetetlenebbé válik az intelligenciá, kreativitás és cselekvőképesség kizárólag embereknek vagy gépeknek való hozzárendelése.

---

7. György Csepeli — közepes-magas eredetiség; legerősebb szintézis

A kötet itt óvatos, és egyetértek ezzel az óvatossággal. A szingularitás, transzhumanizmus, Industry 4.0, megfigyelő kapitalizmus, okosvárosok és szuperintelligencia nem Csepeli találta fel. Ennek ellenére több kombináció is kifejezetten az övé.

A legfontosabb az **episztemológiai–axiológiai megkülönböztetése** , amelyet az MI-re alkalmazzák. A gépi rendszerek fokozatosan átsajátíthatják a tudás szféráját – tényeket, mintákat, előrejelzést, optimalizálást és technikai problémamegoldást –, miközben megoldatlanul hagyják az érték szféráját: **mit kellene kívánni és miért?**

Ez Csepeli legerősebb paradoxonát eredményezi: **az episztemológiai mindenhatóságot az axiológiai ürességgel kombinálva**. Egy civilizáció rendkívül hatékony lehetne az eszközök megtalálásában, miközben nem lesz jobb a célok meghatározásában.

Egy másik megkülönböztető szerkezet a **Robot sapiens, az Ember 2.0 és az Új Herceg közötti** megkülönböztetés. Három logikailag eltérő utat képviselnek:

Robot sapiens = mesterséges versenytárs; Human 2.0 = átalakult/hibrid ember; New Prince = hálózatba kötött kollektív intelligencia, amely infrastrukturális vagy szuverén hatalmat szerzett.

Ez sokkal kifinomultabb, mint a "jövőbeli MI"-t egyetlen homogén végpontként kezelni.

Egy harmadik értékes szintézis ötvözi a Heideggeri **Machenschaftot** az **önkéntes digitális megfigyeléssel**: a technológiailag engedélyezett teljes megfigyelés átfogóbbá válhat éppen azért, mert az egyének önként generálják és adják át azokat az adatokat, amelyeken működik.

A kötet maga öt jellegzetesen csepeli-i konstrukciót azonosít: az adatficációtól hatalomra vezető oksági sorrendet, az episztemológiai/axiológiai megkülönböztetést, a Robot sapiens/Human 2.0/New Prince architektúrát, az önkéntes digitális megfigyelést, amelyet gépészeten keresztül értelmeznek, és a munkanélküliség átalakulását a jövedelmi problémából jelentés problémájává.

Globális értékelés: Ezeket inkább **eredeti szintézisnek nevezném, mint a világelső konceptuális találmány erős** állításainak. Csepeli eredetisége építészeti: számos nemzetközi MI-vitát szervez egy evolúciós elméletté az emberi lény változásáról.

---

8. Ivana Greguric — közepes-magas eredetiség, egy különösen erős koncepcionális újítással

Greuric legerősebb, azonosítható koncepcionális újítása a **kiborgoetika**: nem csupán egy újabb bioetika alkalmazása, hanem egy szisztematikus normatív terület annak meghatározására, hogy hol válik a terápiás technológiai beavatkozás fejlesztéssé, hol a fejlesztés újratervezésé, és hol az újratervezés megszüntetheti azt az emberi alanyt, akit állítólag fejlesztenek.

Az alapvető paradoxon erőteljes:

Az emberi korlátlan fejlődés nem konvergál egy végtelenül fejlődő emberhez; egy ponton túl egy másfajta lényt is létrehozhat.

Második fontos lépése az **eszközetikától az ontológiáig**. A technológia már nem csupán az emberek kidolgozására hat. A kibernetika, az MI, a neurotechnológia és a biotechnológia képes módosítani azokat az entitásokat, amelyek először lépnek be a társadalmi kapcsolatokba.

Harmadszor, két olyan irodalmat integrál, amelyeket gyakran külön-külön tárgyalnak:

külső mesterséges intelligencia + belső emberi fejlesztés → kibernetikui/poszthumán ügynökségi formák.

Negyedik a kötet **antropológiai kivándorlásnak** nevezett helye. A fizikai kihalás nem szükséges az emberiség filozófiai értelemében a "végéhez". A biológiai emberek túlélhetnek, miközben az autonómia, testi integritás, stabil személyes identitás, a gondolat magánélete és az exkluzív emberi történelmi cselekvés eltűnik vagy radikálisan átalakul. Ez a négy hozzájárulás kifejezetten megszűnik a kötetben.

Globális értékelés: a transzhumanizmus, a kiborg-elmélet és a poszthumán ontológia mind nagy nemzetközi hagyományokkal rendelkezik, így az ember-gép konvergencia általános tézise nem globálisan új. A legerősebb igény a független újdonságra a **kiborgoetika mint integrált normatív program**, amelyet követ a külső mesterséges intelligencia és belső emberi fejlesztés szisztematikus egysége.

---

9. Nenad Vertovšek — közepes-magas eredetiség, főként szintézis

Vertovšek legérdekesebb független tézise az **MI kockázatának antropogén fordított** változata. Az elsődleges veszély nem feltétlenül egy etikailag hibás mesterséges intelligencia legyen; lehet, hogy **az emberi intelligencia már beágyazódott a profit, háború, manipuláció és dominancia intézményeibe, amely egy rendkívül erős MI-erősítőt** szerzett.

Tömörített formában:

AI kockázat = gépi képesség × társadalmi struktúra × emberi etikai hiány.

Ez elmozdítja az elemzést a jövő gépének hipotetikus pszichológiájáról, és a jelenlegi és jövőbeli MI-t irányító intézmények felé tereli.

Az emlékezetes megfogalmazása, miszerint **"nem robotok, hanem algoritmusok"** dominálhatnak ugyanabban az irányban. A hatalom nem feltétlenül az autonóm gépi vonalzó antropomorf formáját ölti. Beágyazhat az infrastruktúrákba, amelyek szabályozzák a láthatóságot, ajánlást, osztályozást és hozzáférést.

Egy kapcsolódó hozzájárulás az MI-től az MI-hez mint eszközhöz való átalakulása, **mint az észlelt valóságot alkotó médium**. A hagyományos propaganda manipulálja az üzeneteket; az algoritmikus rendszerek képesek manipulálni azokat a feltételeket, amelyek meghatározzák, mely üzenetek és valóságok válnak láthatóvá.

Vertovšek ezt a **technofeudalizmushoz** köti: rendkívül fejlett technológiai rendszerek együtt létezhetnek egyre hierarchikusabb társadalmi viszonyokkal. Azonban maga a kifejezés nem az ő találmánya; a kötet kifejezetten megemlíti, hogy használta Varoufakis fogalmát. Így az eredetiség abban rejlik, hogy az algoritmikus közvetítést, az MI-t és a poszthumán elméletet összekapcsolja ezzel a politikai-gazdasági diagnózissal, nem magával a "technofeudalizmusban".

A kötet maga a jellegzetes architektúráját hét összekapcsolt állításként rekonstruálja: MI mint a meglévő kapcsolatok erősítője; az emberi intelligencia etikai hiánya; MI mint valóságépítő médium; az infrastruktúra irányítása mint hatalom; technológiailag fejlett, de hierarchikus társadalmi kapcsolatok; az ember–gép határ feloldása; és a legtávolabbi horizonton hibrid/nem emberi intelligencia.

Globális értékelés: az egyes komponenseknek jelentős precedensei vannak. Vertovšek eredetisége tehát elsősorban **szintetikus és relációs** jellegű, különösen a médiaelmélet, az etikai kudarc, az algoritmikus hatalom és a poszthumán átalakulás kombinációja.

Összehasonlító következtetés: mely állítások érdemlik meg a legerősebb globális elméleti újdonságcímkét?

Miután valóban új elméleti mechanizmusokat választunk el a létező nemzetközi eszmék széleskörű szintéziseitől, a kötet saját rangsorolását a következőképpen élesítenék.

Az eredetiség rangja	Szerző	A legerősebb jelölt az igazán egyedi hozzájárulásra
I/A+	**Béla Pokol**	A társadalom mint kollektív szuperintelligencia; autonóm MI mint egy hartmann-i evolúciós restart, amely megkerüli a biológiai/pszichikus rétegeket
I/A+	**Jacek Dukaj**	Önkéntes deszubjektizálás felsőbbrendű optimalizáláson keresztül; a szuverenitás láthatatlan átruházása; Az AI mint a Logos külső megjelenésének csúcspontja
I/A+	**Vladan Joler**	MI mint bolygói anyagkitermelő infrastruktúra; a Föld és az emberi kogníció egyidejű kolonizációja; privatizált osztályozás/előrejelzés
I/A+	**Mihai Nadin**	Radikális MI társadalmi átalakulás valódi gépi intelligencia nélkül; az emberek és intézmények gépszerűvé válnak a számítási alkalmazkodás révén
I/A	**Andrzej Zybertowicz**	Rekurzív digitális előfeltétel + két görbe + techno-csapda: az MI az emberekre hat hat, akiket már kognitívan/intézményileg átalakítottak a korábbi digitalizáció
II/A−	**József Kelemen**	Kölcsönös gépi–ember antropológiai rekonstrukció és az intelligencia/kreativitás hibrid megjelenése
II/B+	**Ivana Greguric**	A kiborgoetika és a külső mesterséges intelligencia integrációja az emberi belső áttervezéssel egy normatív-ontológiai problémává
II/B+	**György Csepeli**	Episztemológiai mindenhatóság kontra axiológiai ürességː Robot sapiens–Human 2.0–New Prince háromszoros architektúra
II/B+	**Nenad Vertovšek**	Az MI veszélyének antropogén megfordítása és az MI mint algoritmikus médium, amely társadalmi valóságot alkot

Ez szorosan megfelel a kötet saját következtetésének, miszerint **Pokol, Dukaj, Nadin és Joler a legmagasabb kategóriát foglalja el a fogalmi eredetiségben**, **Zybertowicz közvetlenül alattuk**, míg Kelemen, Greguric–Vertovšek és Csepeli főként jellegzetes szintezisekkel járulnak hozzá.

Van azonban még érdekesebb következtetés. Ezek az elméletek eredetiek különböző **dimenziókban, ezért nem redukálhatók egyetlen skálákra. Pokol újdonsága elsősorban** szociológiai-ontológiai**; Dukaj** antropológiai-civilizációs**; Joler** materialista/politikai-közgazdasági**; Nadin** szemioi-előrelátó**; és Zybertowicz** intézményi-civilizációja. Más szóval, nem csak abban, hogy az MI jó vagy rossz volt-e, hanem **abban, hogy az MI forradalom alapvetően mit alakít át**: a társadalom kognitív architektúráját, az emberi szubjektivitást, a hatalom infrastruktúráját, az élő intelligencia előrelátó szerkezetét, vagy a civilizáció képességét, hogy irányítsa saját technológiai fejlődését. A kötet kifejezetten ezt az öt alternatívát jelöli meg a legfontosabb összehasonlító eredményként.

A kötet egészére vonatkozó legvédettebb állítás tehát nem az, hogy ezek a szerzők teljesen példátlan fogalmakat találtak ki. Ehelyett **legalább öt valóban egyedi elméleti program jelenik meg, amelyek sokkal nagyobb láthatóságot érdemelnek a világméretű MI és társadalmi irodalomban**: Pokol **társadalmi/ontológiai szuperintelligenciája**, Dukaj **önkéntes posztszubjektivitása**, Joler **bolygókitermelési MI-je**, Nadin **nem intelligens számítástechnikai emberiség-átalakítása**, és Zybertowicz **rekurzív techno-csapdája egy digitálisan előre kondicionált** civilizációról. Ezeket a téziseket hangosabban hangsúlyoznám egy zárófejezetben, amely a kötet nemzetközi MI társadalomelmélethez való hozzájárulását kívánja bemutatni.

****

Ezen erős originalitási követelmény alapján arra kérem, hogy válassza ki a terület más európai országokból származó kutató-szerzőit, és részletesen ismertesse, és elemezze téziseiket az MI társadalomra gyakorolt hatásairól.

A csatolt kötetet mércéként használva a nyugati/északi/déli európai választást **sokkal szűkebbnek tenném, mint egy normál "MI és társadalom" áttekintés.** A köteted koherens MI-fogalmat, több társadalmi szférán átható hatást, azonosítható ok-okozati mechanizmust, radikálisan átalakult jövő felé való kiterjesztést, valamint egy monográfián vagy kapcsolódó munkásságon belüli fejlődést; a koncepcionális eredetiség további kiválasztási kritérium. Ezután kiemeli a legmagasabb eredetiségi szintet a Pokol kollektív-szuperintelligencia/ontológia tézisét, Dukaj külső Logos-ját és posztszubjektivitását, Joler bolygókitermelő infrastruktúráját, valamint Nadin előrelátási/számítási megkülönböztetését, majd szorosan követi Zybertowicz techno-trap és kapcsolódó modelljei.

Ugyanazt a követelményt **alkalmazva**, nem csupán kiemelkedő MI-etikusok felsorolásával, hanem a következő európai elméletalkotókat választanám. Az "eredetiség" itt egy megkülönböztető elméleti konstrukciót jelent a nemzetközi irodalomban; nem szabad úgy értelmezni, hogy sehol sem található elődő.

Szerző	Fő európai helyszín	Legjellegzetesebb tézis	Az eredetiségről alkotott értékelésem
Luciano Floridi	Olaszország / Oxfordi hagyomány	**Az MI mint ügynökség intelligencia nélkül; a társadalom "körülöleli" magát, hogy az MI hatékonysá tegye**	Nagyon magas
Elena Esposito	Olaszország / Németország	**Mesterséges kommunikáció a mesterséges intelligencia helyett; algoritmusok előre jelzett jövőt gyártanak**	Nagyon magas
Antoinette Rouvroy	Belgium	**Algoritmikus kormányzás: kormányzás korrelációkon keresztül, amelyek egyre inkább elkerülik a témát**	Nagyon magas
Bernard Stiegler	Franciaország	**Automatikus társadalom és általánosított proletarizáció: az automatizálás külső formákat emel és pusztít el a tudás formáiban**	Nagyon magas
Nick Bostrom	Svédország / Oxford	**Ortogonalitás, instrumentális konvergencia, döntő stratégiai előny és az AI singleton**	Alapok / nagyon magas
Éric Sadin	Franciaország	**Az MI mint egy hatalmas, amely "kimondja az igazságot" és a végzési hatalom gyakorlására**	Magas–nagyon magas
Helga Nowotny	Ausztria	**A prediktív MI előállítja azt a jövőt, amit előre jelez, és ezzel szűkíti a nyitott jövőt**	Magas–nagyon magas
David Runciman	Egyesült Királyság	**Az államok és vállalatok már most is mesterséges ügynökök; az MI az emberi cselekvés következő "átadása"**	Magas
Mark Coeckelbergh	Belgium / Ausztria	**Az MI "mesterséges hatalom"; legmélyebb demokratikus hatása az episztemikus cselekvés eróziója**	Magas

1. Luciano Floridi: a társadalom MI-szerűvé válhat, mielőtt az emberszerűvé válna

Floridi az egyik legvilágosabb megfelelője a kötetedben szereplő legeredetibb elméleti alkotóknak. Korábbi infoszféraelmélete és **az "onlife" állapot** már elutasította azt az elképzelést, hogy a digitális technológia pusztán egy eszköz, amely egyébként változatlan társadalomban található. Amint az online és offline határ áthatolhatóvá válik, az IKT **környezeti erővé** válik: a munka, oktatás, egészségügy, személyes kapcsolatok, jog, pénzügy, politika és háború egyre inkább informálisan szervezett környezetben zajlik.

Későbbi MI-elmélete radikálisabb fogalmi lépést tesz. Floridi azt állítja, hogy a kortárs MI-t egy **új ügynökségi formának kell értelmezni, amely intelligencia nélkül is sikeres** lehet. A

régi kérdés, "Képes úgy gondolkodni a gép, mint mi?" ezért félrevezetővé válik. Az AI mérnöki fejlesztése rendkívül sikeres volt éppen azért, mert a rendszereknek nem kell emberi kogníciót reprodukálni ahhoz, hogy célcélra irányított feladatokat hatékonyan végrehajtsanak. A kiegészítő mechanizmust Floridi úgynevezett **átburkolásnak** nevezi: az emberek újratervezik a környezeteket, hogy viszonylag nem intelligens mesterséges ügynökök sikeresen működhessenek bennük.

Ez egy jelentős szociálelméleti inverziót eredményez:

Az emberi világ → géppel olvasható környezet → hatékony mesterséges cselekvési tényező → növekvő intézményi függőség a mesterséges cselekvéstől → a környezet további átalakítása az MI követelményei körül.

Egy raktárhoz nem szükséges olyan robot, aki megvan, mint a raktári dolgozó, józan ész értéssel, ha magát a raktárt átszervezik, hogy a robot navigálni tudja rajta. Ugyanez az elv fokozatosan alkalmazható az adminisztrációra, kereskedelemre, szállításra, kommunikációra és döntéshozatali eljárásokra. A gép nem feltétlenül válik emberré; **a környezet egyre inkább kompatibilissé válik a géppel**. Floridi kifejezetten figyelmeztet, hogy a beburkolás azzal végződhet, hogy magukat az embereket is ösztönözik, hogy alkalmazkodjanak az okos technológiákhoz.

Ez egy szinten nagyon közel áll Nadinhoz, máshol élesen eltérő. Nadin azt állítja, hogy a sikeres számítás nem feltétlenül jelenti az élő intelligenciát vagy valódi előrelátást; ennek ellenére a társadalom átszervezheti magát körülötte. A csatolt kötet tehát fogalmazza meg Nadin radikális lehetőségét, hogy az emberek gépszerűvé válnak, nem pedig gépek emberszerűvé. Floridi hasonló célt ér el inkább **az ügynökség révén, mint előrejelzéssel**: egy rendszer nem kell intelligensnek lennie emberi ontológiai értelemben, ha a környező világot áttervezték, hogy a mesterséges ügynöksége működjön.

Ezért Floridi **ügynökségi/hírszerzési válását + átfogó tézisét** a legeredetibb európai hozzájárulások közé sorolnám. Vitathatatlanul ez a legjobb nyugat-európai megfelelője a Nadin–Joler antiantropomorf vonalának a kötetedben.

---

2. Elena Esposito: az MI forradalom valójában a kommunikáció forradalmának

Esposito elmélete rendkívül fontos az általános szociológia szempontjából, különösen azért, mert kifejezetten a rendszerelméletre épít. 2022-ben megjelent *"Mesterséges kommunikáció: Hogyan teremtenek algoritmusok társadalmi intelligenciát" című* könyve azt javasolja, hogy hagyjuk fel a megszokott kérdést, hogy valóban gondolkodnak-e a gépek, és inkább azt kérdezzük, hogy részt vehetnek-e a **kommunikációban**. Válasza igen: ha az algoritmusok hozzájárulnak a társadalmi intelligenciához, az nem feltétlenül azért van, mert megtanultak úgy gondolkodni, mint az emberek, hanem mert **az emberek és a társadalom megtanultak algoritmusokkal kommunikálni**.

A fogalmi innováció jelentős. A társadalmi relevanciához nem szükséges gépi tudatosság. Egy algoritmus **interakciópartnerré** válhat akkor is, ha nincs szubjektív megértése arról, amit közvetít. Esposito szerint a gépi tanulás mesterségesen reprodukálja a kommunikációs kompetenciát az emberi intelligenciával szemben, lehetővé téve a társas kommunikációban való oksági részvételt, miközben kognitívan eltér az emberi elmétől.

Ez megváltoztatja az elemzési egységet. Ahelyett, hogy

Emberi intelligencia ↔ gépi intelligencia

a releváns reláció a következőképpen

Emberi kommunikációs ↔ algoritmikus működés ↔ Utólag emberi kommunikáció.

Miután egy algoritmus kiválasztja a híreket, mondatot javasol, rangsorolja a jelentkezőket, zenét ajánl, kérdésekre válaszol vagy választ generál, a kimenet belép a társadalmi kommunikáció láncába. Szociológiailag már nem az számít, hogy a gép "valóban megértette-e" az üzenetet. Más résztvevők reagálnak rá, így társadalmilag működőképes lett.

Egy másik, nagyon eredeti összetevő Esposito **személyre szabásról szóló** elemzése. A hagyományos tömegkommunikáció nagyjából ugyanazt az üzenetet küldi mindenkinek. Az algoritmikus kommunikáció lehetővé teszi a tömeges léptéket és az egyéni kezelés sajátos kombinációját: emberek milliói kapnak egyszerre különböző hírfolyamokat, hirdetéseket, ajánlásokat és javaslatokat. A médium a gyártás során személytelen marad, mégis minden címbevevő személyre szabott kommunikációs környezettel találkozik.

Az előrejelzési elmélete még mélyebbre nyúlik. Esposito azt állítja, hogy az algoritmikus előrejelzés nem egyszerűen fedez fel egy már létező jövőt. Egy ajánlás vagy előrejelző osztályozás beavatkozik a jelenbe, és így segíthet **a megjósolt jövő megalkotásában**. Ha egy algoritmus előre jelzi, hogy szeretnék egy terméket, és ennek következtében megmutatja nekem, a későbbi vásárlásom részben megerősíti azt a jövőt, amely nem feltétlenül létezett volna előrejelzés nélkül. A prediktív rendőrségben, biztosításban, orvostudományban és hasonló területeken ugyanaz a rekurzív mechanizmus intézményesen következményekkel válik.

Az így létrejött hurok:

A korábbi adatok → előrejelzést → jelenlegi beavatkozást → megváltoztatott viselkedést → megerősítések/új adatok → javított előrejelzést.

Ez az egyik legerősebb, valóban szociológiai MI-elmélet jelenleg elérhető. A Bielefeld PREDICT program ennek megfelelően pontosan ilyen problémákat vizsgál a biztosításban, precíziós orvoslásban és prediktív rendőrségben, az algoritmikus előrejelzést a társadalom bizonytalansághoz való viszonyának átalakulásaként kezelve.

Esposito még azt is állítja, hogy a múltbeli adatokból való túlzott tanulás veszélyeztetheti a **jövő nyitottságát**: a jóslás szűkítheti azokat a lehetőségeket, amelyekkel az intézmények és az egyének szembesülhetnek.

A köteted szerzőihez képest Esposito egyszerre **áll legközelebb Pokolhoz, Nadinhoz és Jolerhez**, de egyikükre sem redukálható. Pokol az intelligencia egységét az egyéni gépről a társadalmilag integrált kollektív kognitió felé mozdi; Nadin elválasztja az előrejelzést az élő várakozástól; Joler hangsúlyozza a besorolást és az előrejelzést, mint a hatalom infrastruktúráját. Esposito egy kifejezetten luhmann-i alternatívát kínál: **az MI társadalmilag forradalmi lesz, mert a kommunikáció részben nem emberivé válhat anélkül, hogy a társadalomnak várnia kellene a valóban emberszerű gépi intelligenciára.**

A koncepcionális eredetiség tekintetében a **legmagasabb európai kategóriába** sorolnám.

---

3. Antoinette Rouvroy: algoritmikus kormányzás és kormányzás, amely egyre inkább elkerüli a témát

Antoinette Rouvroy belga elméleti alkotó az egyik legbefolyásosabb nemzetközi fogalmi újítást hozta létre ezen a területen. Thomas Berns-szel együtt megfogalmazta az **"algoritmikus kormányzást".** Későbbi munkája ezt egy sokkal nagyobb Big Data-elméletté, profilozás, tudás és szubjektíváció elméletévé fejlesztette. Rouvroy saját intézményi leírása szerint kutatásai ötvözik a szemiotikai-epistemikus kérdéseket, a hatalmat és az individuációt, hogy megvizsgálja, miként változtatják meg a számítási rendszerek azt, ami valóságnak számít, és hogyan irányítják az egyéneket és a kollektívákat.

A klasszikus kormányzás az alanyokat kezeli. Törvényeket, normákat, parancsokat, ideálokat vagy átlagokat állapít meg, és megmondja az embereknek, hogyan várják el tőlük a viselkedést. Az algoritmikus kormányzás a Rouvroy–Berns formulációban valami szerkezetileg más irányba halad: **valós idejű, immanens statisztikai normativitás, amely az adatok közötti viszonyokból származik**, és amely hajlamos elkerülni a közvetlen konfrontációt az egyéni alannyal. Az eredmény paradox módon csökkentheti magát a szubjektizálás lehetőségeit.

A sorozat nagyjából következik:

Az egyéni viselkedés → adatok → relációs korrelációkat → profilt/valószínűséget → előrejelző beavatkozást → módosított cselekvési területet követnek.

Az illetőt nem kell meggyőzni. Valójában a rendszernek nem kell tudnia, miért cselekszik. Lehet, hogy elég, hogy bizonyos adatminták korrelálnak a hitelkockázattal, a vásárlási valószínűséggel, a bűnözési kockázattal, a betegség valószínűségével vagy az üzenetre való érzékenységgel.

Ez egy mélyreható átmenet a **kormányból a jelentésen** keresztül **a korreláción keresztüli kormányzatig**.

A radikális pont az, hogy az algoritmikus kormányzás működhet *a* kifejezett szubjektív mérlegelés alatt vagy előtte. Ahelyett, hogy azt mondaná, "Így kellene viselkedni", módosíthatja az árakat, rangsorokat, láthatóságot, lehetőségeket, ajánlásokat vagy hozzáférést, hogy maga a viselkedési környezet is megváltozzon.

Rouvroy újabb **"algoritmikus realizmus" fogalma** még erősebbé teszi ezt az állítást. 2025-ös megfogalmazásában az MI által feldolgozott hatalmas adatok egyre inkább nemcsak a valóság ábrázolásaiként, hanem kiváltságos koordinátákként is működnek, amelyeken keresztül maga a valóság keletkezik. Politikai szinten ugyanez a "realizmus" az MI alkalmazását kívánatosnak, sürgősnek vagy elkerülhetetlennek mutatja be, ezzel lezárva a közképzeletet és a demokratikus vitákat.

Az elmélet tehát három történelmi szakaszon halad át:

normatív kormányzat → statisztikai kormányzat → algoritmikus környezeti kormányzat.

A harmadik szakaszban a hatalomnak nem kell megvárnia, hogy egy autonóm személy döntse el, engedelmesked-e neki.

Ez különösen közel tartja Rouvroyt Jolerhez. Joler az MI-t a besorolás és előrejelzés privatizációjaként és automatizálásaként írja le; a köteted ezt az egyik legjelentősebb politikai meglátásaként kezeli. Rouvroy kiegészítő elméletet ad arról, **milyen kormányzati racionalitás alakul ki, amikor maguk a besorolás és a jóslat válik a hatalom közvetítőjévé**.

Ő is keresztezi Dukaj-t: Dukaj problémája az önkéntes delegáció által elhalványuló autonóm alany; Rouvroyé más értelemben radikálisabb – a kormányzati gépezet egyre inkább feleslegesnek tartja, hogy a témát reflektor alanyként kezelje.

Az algoritmikus kormányzást **kétségtelenül első rangú európai elméleti hozzájárulásként sorolnám be.**

---

4. Bernard Stiegler: a technológiai automatizálástól az emberi tudás proletarizálódásáig

Bernard Stiegler egy kis minősítést igényel. Elmélete megelőzi a kortárs generatív MI hullámot, és a digitális automatizálás, algoritmusok és az "automatikus társadalom" szélesebb kategóriáit használja. De éppen ezért jogosult a rendszerépítési kritériumba: az MI egy sokkal régebbi, az emberi memória, tudás és cselekvőképesség technikai külső megjelenésének filozófiájába épül.

Kulcsfontosságú fogalma a **proletarizáció**, de ezt messze túlmutat a szegény bérmunka marxista jelentésén. Stiegler szerint a technikai rendszerek képesek olyan külső tudásformákat kiszélsíteni, hogy az emberek fokozatosan elveszítik a képességüket arra, hogy elvégezzék vagy megértsék, mit végeznek nekik a gépek. Megkülönbözteti a **svoir-faire** – a dolgok megteszésének tudása; **savoir-vivre** – az élet ismeretét; és **a svoir-théoriser** – a koncepció és elméletalkotás képességét. A kortárs Stiegler-kutatás pontosan ezt a tudásközpontú definíciót hangsúlyozza.

Ez az MI automatizálásának sokkal mélyebb jelentést ad, mint a munkahelyhelyettesítés:

Az emberi tudás → formalizációja → technikai exteriorizáció → automatizálás, → a tudás belső elsajátításának csökkenése → az automatizált rendszertől való függőség.

Egy munkás elveszítheti a gyakorlati készségeit. Egy fogyasztó elveszítheti az autonóm életmódját. Végül még a szakértők és döntéshozók is elveszíthetik az elméleti megértést, ha a számítás olyan válaszokat ad, amelyeket már egyetlen embernek sem kell rekonstruálnia.

Az Automatikus társadalom: A munka jövője *című művében* Stiegler összekapcsolja ezt az episztemikus mechanizmust a politikai gazdaságtannal. Az adatgazdaság által generált integrált automatizálás fokozatosan megszüntetheti a foglalkoztatást mind a manuális, mind a szakmai területeken. Normatív válasza tehát nem csupán az egyetemes jövedelem. Megkülönbözteti a **bérért végzett foglalkoztatást** a **gazdagabb tudásos, értelmes részvétel értelmű munkától, és olyan járulltató gazdaságot javasol, amely a tudás és részvétel fejlesztésére irányul, nem pedig az embereket passzív haszonélvezőivé tenni az automatizált termelésből.**

Ez kulcsfontosságú. A hagyományos automatizálási kérdés a következő:

Milyen jövedelmet kapnak az emberek, ha a gépek végzik a munkájukat?

A Stiegler-kérdés a következő:

Mi történik a civilizációval, amikor azok a tevékenységek, amelyek révén az emberek kompetenciát, autonómiát, ítélőképességet és társadalmi részvételt szerzetenek, már nem végzik maguk által?

Válasza az **általánosított kognitív proletárizáció** lehetősége. Az MI technikailag képesebbé teheti a társadalmat, miközben emberi tagjai fokozatosan kevésbé képesek arra, hogy képesek reprodukálni azt a tudást, amelyre a társadalom támaszkodik.

Stiegler ellenfogalma a **deautomatizáció**. Az automatizálás nem egyszerűen tiltható – a technikai rendszer gyógyszer, egyszerre mérgező és gyógyító. A probléma az, hogy technikai és társadalmi megállapodásokat építsünk, amelyek lehetőséget adnak az értelmezésre, találmányra, tanulásra és kollektív individualizációra az embereknek. Későbbi munkái összekapcsolják ezt a negentrópiával: az élő tudás újdonságot és kettéosztást teremt, míg a teljes algoritmikus automatizáció az ismétlődés és entrópia felé hajlik.

Stiegler tehát feltűnően közel **áll Dukajhoz**, de a különbség pontos. Dukaj fenyegetett entitása az **alany**: a külső Logosz önálló ítélkezést kulturálisan feleslegessé teszi. Stiegler fenyegetett entitása a **tudás, amely személyekben és kollektívákban testesül meg.** Mindkettő a kiesőjelet írja le, de Dukaj posztszubjektivitással zárul, míg Stiegler általánosított proletarizációval.

Erős összefüggés van Pokol technológiai munkanélküliségről alkotott elméletével is. Pokol azt kérdezi, hogy a hatalmas automatizálás végső soron megszüntetheti-e a piacok funkcionális szükségletének egy részét, és ezzel növelheti a politikai-demokratikus ellenőrzés fontosságát; a csatolt kötet ezt az egyik jellegzetes kiterjesztésének nevezi. Stiegler ehelyett azt kérdezi, hogyan őrizhetné meg az emberi tudást és az individuációt a munkaviszony utáni rendszer.

Ezért Stieglert a **nagyon magas eredetiségű kategóriába** sorolnám, annak ellenére, hogy a terminológiája "automatizálás", nem pedig a mai LLM-központú MI.

5. Nick Bostrom: Az MI megváltoztatja a globális hatalom szerkezetét, nem csupán a termelést

Bostrom már az egyik fő összehasonlítási pont Pokol a kötetedben, de nyugat-európai kiterjesztésbe tartozik, mert oxfordi munkája az egyik alapvető európai rendszer a radikális MI gondolkodásához.

Első kulcsfontosságú eredeti állítása az **ortogonalitás tézis**: az intelligencia szintje és a végső célok nem kell, hogy változzanak egymással. Egy radikálisan intelligensebb rendszer így nem válik erkölcsileg bölcsebbé, és nem konvergál automatikusan az emberi értékekhez. A második az **instrumentális konvergencia**: nagyon különböző végső célok hasonló köztes ösztönzőket generálhatnak, mert önvédelem, erőforrások megszerzése vagy hatékony irányítás növelése széles körű végső célokat eredményezhet.

Ez lerombolja az egyik legmegnyugtatóbb intuíciót a szuperintelligenciáról:

Nagyobb intelligencia ≠ nagyobb erkölcs.

A társadalmi következmények akkor kezdődnek, amikor egy rendszer eléri azt, amit Bostrom döntő **stratégiai előnynek nevez**. *A szuperintelligencia* tehát túlmutat a technikai MI-n az intelligencia robbanásának, kognitív szuperképességeknek, multipolaritásnak, értékszerzésnek és a kialakuló világ stratégiai struktúrájának kérdéseire.

Korábbi fogalma az **egyesről** szociológiailag különösen jelentős. Az egy-egy világrend, amely egyetlen hatékony döntéshozó szervet tartalmaz a legmagasabb szinten, amely elég erős ahhoz, hogy fenntartsa felsőbbrendűségét és hatékony ellenőrzést gyakoroljon a terület főbb jellemzői felett. Nem kell diktatúra legyen: egy demokratikus világrendszer elvileg lehet az, de egy elég erős mesterséges szuperintelligencia is.

Így Bostrom sorozata nem csupán az:

A gép okosabb lesz az embernél.

Ez potenciálisan így van:

A szuperintelligencia döntő stratégiai előnyt → → a versengő hatalmi központok megszüntetésében, → Singleton → hosszú távú stabilizációja vagy egy fejlődési pálya "bezárása".

Így az MI az **egész civilizáció alkotmányos rendjének** problémájává válik.

Az Armstrong–Bostrom–Shulman "a szakadék felé való versenyzés" modellje hozzáadja azt az intézményi mechanizmust, amely összekapcsolja a technológiai helyzetet a társadalmi versengéssel. Ha több fejlesztőcsapat úgy véli, hogy az első szerep óriási előnyt teremt, az

egyénileg racionális versenyzők csökkenthetik a biztonsági kiadásokat, hogy elkerüljék a verseny elvesztését. Több versenytárs és nagyobb ellenséges tényező növelheti a rendszerszintű veszélyt.

Ez különösen feltáróvá teszi a **Zybertowicz-szal való** összehasonlítást. Zybertowicz techno-csapdája szerint a piaci és geopolitikai verseny egyénileg irracionálissá teszi a korlátozást, még akkor is, ha a gyorsulás kollektíven veszélyes. A csatolt kötet helyesen azonosítja ezt az intézményi problémaváltozatot, mint az egyik legeredetibb szociológiai hozzájárulását. Bostrom formalizált egy speciális MI-faj változatot; Zybertowicz ezt a mechanizmust egy sokkal tágabb civilizációs szociológiába ágyazza.

A Pokollal **való kontraszt** még élesebb. Bostrom elsődleges szakadása egy rendkívül képes **mesterséges ügynök** megjelenése. Pokol elmélete az elemzés egységét a társadalom kognitív integrációjához helyezi: kollektív szuperintelligencia kialakulhat emberek és gépek hálózatán keresztül, akár egyetlen autonóm szuperelme előtt, akár nélkül. Köteted kifejezetten azonosítja ezt az elemző egység elmozdulását Pokol egyedi újításaként.

Így Bostrom továbbra is alapvető és rendkívül eredeti marad, de az összehasonlítás azt is mutatja, hogy a kelet-európai elméletek szociológiailag hol lépnek túl rajta.

---

6. Éric Sadin: mesterséges intelligencia tekintélyként, amely "kimondja az igazságot"

Éric Sadin elmélete az egyik legerősebb francia megfelelője Dukajnak és Rouvroynak. Ugyanabba a széles körű "elméleti" értelmezésbe vonom őt, amely lehetővé teszi a kötetedben, hogy Dukaj is szerepeljen: filozófus és esszéista, nem hagyományos empirikus MI-kutató.

A La Vie algorithmique *(2015) című művében* Sadin már úgy írja le a digitalizációt, mint egy olyan világot, amelyben az érzékelők által generált adatok a **valóság érthetőségének kiváltságos forrásává** válnak. Egyre gyakrabban szervezik a kereskedelem, az oktatás, az orvostudomány, az interperszonális kapcsolatok, a városi élet és a háztartási környezetek valós idejű numerikus értékelés és optimalizálás által.

2018-as *L'Intelligence artificielle ou l'enjeu du siècle című művében* a döntő elméleti lépést teszi meg. Sadin szerint az MI jellegzetes társadalmi funkciója nem alapvetően "gondolkodás", hanem az, hogy megszerezze a hatalmat, hogy **kimondja, mi igaz a helyzetekben** – hogy megbízhatóbb valóságértékelést igényeljen, mint az emberi ítélkezés. Ebből az episztemikus tekintélyből következik az, amit ő **injunktív hatalomnak nevez**: a számítási rendszerek nem csupán leírják a helyzeteket; kimenetük egyre inkább megmondja az embereknek és szervezeteknek, mit kell tenniük.

A sorrend a következő:

adatfikáció → számítási értékelés → a felsőbbrendű igazság → ajánlás → előírás feltételezése, → az ítélet → viselkedési orientáció delegálása.

Sadin központi veszélye tehát nem egy tudatos MI lázadás. Az az, hogy az emberi civilizáció önként új forrást teremt az **episztemikus tekintélyben** , és fokozatosan gyakorlatilag kötelezőként kezeli eredményeit.

A lehetséges eredményt egyfajta **automatizált, láthatatlan kézként** írja le: a valóság folyamatos elemzése közvetlenül kapcsolódik a kereskedelmi, adminisztratív és viselkedési beavatkozásokhoz.

Ez meglepően közel áll Dukaj önkéntes deszubjektizálásához, de a mechanizmusok eltérnek. Dukaj azt mondja: az algoritmusok jobb döntéshozókká válnak, így az emberek racionálisan delegálják a döntéseket, amíg a szubjektivitás el nem múlik. Sadin azt mondja: az algoritmusok az **igazságkimondó hatóságok** társadalmi státuszát szerzik, így az ajánlások fokozatosan végzési erőt vesznek át.

Rouvroy ezután egy harmadik variációt ad: az algoritmusnak feltétlenül nem kell a témát egyáltalán kezelnie.

Ezért ez a három elmélet sorozatként rendezhető:

Sadin: számítási hatalom az ítélkezés felett.
Dukaj: önkéntes átruházás az ügynökség felett.
Rouvroy: a kormányzás egyre inkább megkerüli magát az alanyt.

Ez egy rendkívül termékeny elméleti klaszter.

Sadin korlátai módszertan. Elemzése filozófiailag erőteljes és gyakran előrelátó, de retorikája erősen normatív, és kevésbé ajánl empirikus intézményi megkülönböztetést, mint Pokol, Zybertowicz vagy Joler. Ennek ellenére az **igazság/injunktív-hatóság tézist** nemzetközi eredetiségben magas-nagyon magasnak értékelném.

---

7. Helga Nowotny: a prediktív irányítás paradoxonája

Az osztrák tudományos és technológiai elméleti szakember, Helga Nowotny egészen más elméletet dolgoz ki. Az *In AI We Trust: Power, Illusion and Control of Predictive* Algorithms (2021) **című műve az MI társadalmi hatásainak** középpontjába helyezi az időt és a jövőt.

Az ő alapvető paradoxonája egyszerű, de elméletileg erőteljes. Az emberek azért alkalmazzák a prediktív MI-t, mert nagyobb kontrollt akarnak a bizonytalanság **felett. De a jóslás performatív: ha az emberek és intézmények hisznek egy előrejelzésnek, akkor megváltoztatják viselkedésüket válaszul. Így a jóslat elkezdi segíteni azt a valóságot létrehozni, amelyet csak előre állított. Ennek következtében minél inkább támaszkodunk az MI-re a jövő irányításában, annál inkább csökkenthetjük saját** cselekvésünket a jövő felett.

Az oksági hurok a következő:

A múltbeli adatok → előrejelzést → a jóslásban → viselkedés megváltoztatását → önmagát megerősítő jövőt → nagyobb bizalommal a jóslásban.

A szélsőséges veszély nem csupán a pontatlan előrejelzés. Ez **a jövő kulturális jelentésének átalakulása**.

A modernitás fokozatosan nyitottnak fogta fel a jövőt: politikai cselekvéssel, tudománysal, találékonysággal és választással változtatható meg. Egy társadalom, amelyet rendkívül hiteles előrejelző rendszerekkel telítettek, visszasodródhat valami determinizmushoz hasonló irányba. A jövő már megjelenni kezd a modellben belül.

Ezért figyelmezteti Nowotny, hogy a prediktív fordulat bezárhatja a jövő "nyitott horizontját", és a határon az embereket magukat is előrejelzést követő rendszerekké alakíthatja.

Normatív válasza nem csupán "jobb algoritmusok". A bizonytalanság elfogadását, bölcsesség és azt, amit **ő "katedrális-gondolkodásnak" nevez,** olyan intézményi képességeket követel, amelyek képesek összekapcsolni a azonnali cselekvést hosszú távú horizontokkal, amelyeket nem lehet optimalizálási kritériumokra redukálni.

Az elméleti közelség **Espositóhoz** feltűnő. Mindketten azt állítják, hogy az algoritmikus előrejelzés részt vesz a jövő létrehozásában, amelyet megjósol. De Esposito ezt egy kommunikációs és társadalmi rendszerek elméletén belül fogalmazza meg; Nowotny ezt az **idő, az ügynökség és a modernitás bizonytalansághoz való viszonyának** elméleteként fejleszti ki.

Van egy értékes kontraszt is **Nadinnal** szemben. Nadin megkülönbözteti az igazi élő várakozást a számítási előrejelzéstől. Nowotny azt kérdezi, mi történik társadalmilag, még akkor is, ha az előrejelzés nem más, mint előrejelzés: ha az emberek ennek megfelelően szervezik a cselekvést, a jósló gépek átalakíthatják az emberi várakozást.

Ez egy valóban eredeti tézis, bár szűkebb, mint Floridi vagy Stiegler általános rendszerei. Én Nowotny-t közvetlenül a legmagasabb csoport alá helyezném.

---

8. David Runciman: Az MI nem az első mesterséges ügynök, akinek az emberek átadták a hatalmukat

David Runciman brit politikai elméleti szakember az egyik legszokatlanabb közelmúltbeli európai rekonstrukciót dolgozza ki a *The Handover: How We Gave Control of Our Lives to Corporations, States and* MIs című művében.

Alapvető állítása szándékosan provokatív: a mesterséges ügynökség megjelenése **történelmileg nem új**. Az emberek már rendkívül erőteljes, nem biológiai ügynököket, úgynevezett **államokat és vállalatokat építettek** fel. Ezek az entitások sok ember cselekedeteit egyesítik, jogilag elismert identitással rendelkeznek, túlélik tagjaik halálát, erőforrásokat halmoznak fel és olyan képességeket gyakorolnak, amelyeket egyetlen ember sem tudna

egyedül gyakorolni. Ebben a funkcionális értelemben Runciman szerint az emberiség évszázadokkal a digitális MI előtt létrehozott hatalmas "robotokat".

A kortárs MI jelentőségét tehát egy háromszázéves történeten keresztül értik, amely a **mesterséges konstrukciók közötti cselekvési lehetőséget** adta:

Az emberi egyének → mesterséges jogi/intézményi személyek, → államok és vállalatok, akik szuper-egyéniséggel → digitális mesterséges ügynökökkel rendelkeznek→ intézményi és gépi ügynökség lehetséges kombinációja.

A legmélyebb jövőbeli kockázat tehát nem az MI-től, sem a vállalatoktól vagy államoktól elszigetelve származhat, hanem a **kapcsolatukból**.

Egy határvidéki MI, amelyet egy hétköznapi egyén irányít, egy szintű hatalommal rendelkezik. Ugyanaz a kognitív rendszer, amely egy multinacionális vállalatba, katonai bürokráciába vagy hatalmas államba ágyazódik, megszerzi az intézményi ügynök képességeit, amely körülötte van.

Ez szokatlan elméleti perspektívát teremt az MI "átvételéről". Lehet, hogy nincs tiszta esemény, amikor az emberiség átadná a politikai hatalmat a gépeknek, mert az emberek már évszázadokon át építettek intézményeket, amelyeknek átadják az egyéni cselekvési területek részeit.

Runciman áll legközelebb **Dukajhoz**. Dukaj a nyelvet, írást, bürokráciát és MI-t a Logos és az ügynökség külső megjelenésének szakaszaiként értelmezi. Runciman elszigetelte ennek a történelemnek az egyik kulcsfontosságú szálát: **az intézményi mesterséges** ügynökséget.

A különbség az, hogy Dukaj végpontja antropológiai – az autonóm alany, mint a civilizáció központjának eltűnése. Runciman politikai: **milyen irányítást tarthatnak fenn az egyes emberek olyan mesterséges ügynökök felett, akiknek képességei és időtartama meghaladja alkotóikét?**

Ezt nagyon eredetinek nevezném, bár kevésbé kidolgozott MI-társadalom rendszerként, mint a Floridi, Esposito vagy Stiegler.

---

9. Mark Coeckelbergh: a mesterséges intelligencia is mesterséges hatalom

Mark Coeckelbergh belga filozófus, aki a bécsi egyetemen dolgozik, fokozatosan kiépített egy széles körű politikai elméletet az MI-ről. Az *MI politikai filozófiája együtt* kezeli a szabadságot, egyenlőséget, diszkriminációt, demokráciát, megfigyelést, környezeti politikát és poszthumanizmust, és azt állítja, hogy az MI-t nem csupán mesterséges intelligenciáként, hanem **"mesterséges hatalomként" kell érteni.**

A kifejezést nem szabad félreérteni azzal, hogy a gép önállóan birtokolja a politikai hatalmat. Későbbi megfogalmazása kifejezetten relációs jellegű: a hatalmat mesterséges intelligencián

keresztül gyakorolják, amit ő emberi **technoperformansoknak nevez.** Az emberek, intézmények és gépek együtt olyan rendszereket alkotnak, amelyek révén az embereket besorolják, befolyásolják, fegyelmezik vagy felhatalmazzák.

Legelméletileg legjellegzetesebb közelmúlt hozzájárulása az **episztemikus cselekvőséggel** kapcsolatos. Egy működő demokráciához olyan állampolgárok szükségesek, akik nemcsak információnak vannak kitéve, hanem képesek hiedelmet alkotni, bizonyítékokat értékelni és véleményüket felülvizsgálni. Az MI-alapú szűrés, személyre szabás, statisztikai alapértelmezett tudás és episztemikus buborékok növelhetik az elérhető információ mennyiségét, miközben gyengíthetik az emberek bizalmát saját ismeretelméleti képességeikbe és azok gyakorlását.

Az eredmény az oksági sorozat:

Az MI által közvetített információs környezet → algoritmikus kiválasztódástól függő → a megváltozott hiedelmek kialakulása/felülvizsgálata, → gyengült episztemikus cselekvési képesség → gyengült politikai cselekvési lehetőség → demokratikus sebezhetőség.

Ez mélyebb, mint egy "álhír" elmélete. Az állampolgár továbbra is szabadon szavazhat, miközben elveszíti a kognitív autonómiát, amely a jelentős demokratikus állampolgársághoz szükséges.

2024-es könyve, *a Why AI Undermines Democracy and What To Do About It* elemzést kiterjeszt a manipulációra, megfigyelésre, polarizációra, szabadságra és egyenlőségre, miközben a demokratikus technológiák, az intézményi reform és az oktatás mellett érvel, nem pedig technológiai elutasítást.

Coeckelbergh kissé kevésbé egyedi, mint Esposito vagy Rouvroy: összetevői közül sok a meglévő politikai filozófiából származik. Egyedi eredménye az, hogy ezek a **mesterséges hatalom + episztemikus cseleensség** köré integrálják. Ezért a szigorú csoportban egy erős második helyre helyezném, nem pedig a négy-öt legradikálisabb fogalmi újító közé.

Szerzők, akik éppen a szigorú küszöb alatt vannak

Több további európai vagy európai elméleti szakember is nagyon fontos, de a köteted szabványa alapján megkülönböztetném őket a fenti kilencektől.

Mireille Hildebrandt (Belgium/Hollandia) rendkívül eredeti a jogelméletben. *Az Okos technológiák és a jog vége* szerint a társadalom az információs társadalomból egy **adatvezérelt társadalom felé halad** , ahol a gépi tanulás előrejelzi és megelőzi a tevékenységeket, preferenciákat, egészségügyi kockázatokat, hitelképességeket és bűnözési kockázatokat. Kidolgozza az **adatvezérelt ügynökség**, az okos környezetek és végső soron a "tervezés által védett jogvédelem" elképzeléseit: ha maguk a számítási architektúrák egyre inkább cselekvést strukturálnak, a jogi védelemnek részben technológiailag megtestesülnie kell, nem csak nyomtatott szabályokban kell maradnia. Nagyon magasra rangsorolnám őt eredetiség miatt, de

a rendszere továbbra is a jogra, az alkotmányos védelemre és az ügynökségre épül, nem pedig az egész MI társadalom teljesen differenciált elmélete.

John Danaher (Írország) bevezette az **"algokrácia fenyegetése" befolyásos** gondolatát: az algoritmikus közvélemény-döntéshozatal aláássa a politikai legitimitást, mert az emberek egyre inkább olyan döntési eljárásoknak lesznek kitéve, amelyekben az értelmes emberi részvétel és kontroll lehetőségei csökkentik. Az *Automation and Utopia* című műve ezután szinte az MI automatizálás ellentéte – egy kötelező foglalkoztatás nélküli társadalom megfelelő intézmények mellett bővítheti az emberi virágzás lehetőségeit. Mindkét tézis, de a két irányzat még nem alkot olyan átfogó MI-társadalom rendszert, mint Floridié vagy Stiegleré.

Daniel Innerarity (Spanyolország) talán a legerősebb, kifejezetten demokratikus anti-helyettesítő érvet dolgozta ki. Központi tézise nem csupán az, hogy a jelenlegi MI nem elég fejlett a demokratikus kormányzáshoz. A politika természeténél fogva **összetettséggel, bizonytalansággal, kétértelműséggel és véletlenszerűséggel** foglalkozik; az algoritmikus pontosság nem megfelelővé válik, amikor vitatott társadalmi célokról hozott döntéseket kolonizál. Legmélyebb megfogalmazása az, hogy a demokrácia fontos ismeretelméleti értelemben azért létezik, **mert nem tud**: minden rendelkezésre álló bizonyíték figyelembevétele után a valóban politikai kérdések még mindig vitatható döntést igényelnek. Ez nagyon eredeti politikai elmélet, de elsődleges célja a demokrácia, nem pedig a társadalom egészé.

Dan McQuillan (Egyesült Királyság) kifejezetten kritikus elméletet dolgoz ki az *Resisting MI című művében*. Az MI-t úgy határozza meg, mint a mélytanulási műveletek és az őket körülvevő intézmények és ideológiák összekapcsolását, és azt állítja, hogy annak integrációja a megszorításra orientált adminisztratív rendszerekbe fokozhatja a strukturális erőszakot, és algoritmikus "kivételi állapotokat" eredményezhet. Ellenprogramja az absztrakt optimalizálást a gondoskodás, a közösségi javak, kölcsönös segélynek és népi tanácsok váltja fel. Ez egy koherens társadalomelmélet, de erősen antifasiszta normatív kerete miatt általános magyarázó rendszerként kevésbé kiegyensúlyozott, mint Joler szerkezetileg kapcsolódó extrektív elmélete.

Végül **Shannon Vallor**, bár amerikai születésű, megérdemli említést, mint fontos skóciai európai akadémiai hangot. *Az AI Mirror* azt állítja, hogy a mai MI nem annyira világít meg példátlan jövőt, hanem rendkívül erős tükröket **épít az emberiség digitalizált múltjáról**. A modellek felhalmozódott mintákat, előítéleteket és bölcsesség kudarcait reprodukálják, így gátolhatják az emberi képességet az igazán új jövők elképzelésére. Későbbi munkái azt állítják, hogy a történetmesélés és a tudás kiszervezése gyengítheti az emberi narráció, bölcsesség, tudás és cél közötti kapcsolatokat. Fogalmilag ez meglepően közel áll Joler "múlt diktatúrájához", bár Vallor elsősorban erényeetikával és gyakorlati bölcsességgel fejleszti ki a problémát.

A legfontosabb összehasonlító eredmény

Ha ezeket a nyugat-európai elméleteket a kelet-közép-európai rendszerek mellé helyezzük a kötetedben, az eredmény nem kelet–nyugat-ellenállás. Ehelyett egy **transeurópai elméleti családok** halmaza válik láthatóvá.

Az antiantropomorf család magában foglalja Nadint, Floridit és Espositót. Nadin szerint a számítási előrejelzést nem szabad összekeverni az élő előrejelző intelligenciával. Floridi egy lépést félretéli: az igazi intelligencia talán felesleges, mert a nem intelligens mesterséges

ügynökség már képes átalakítani a világot. Esposito a szociológiai lépést teszi: még az intelligencia és az erőteljes cselekvés is másodlagos lehet, ha az algoritmusok már be tudnak lépni kommunikációs láncokba.

A menet rendkívül elegáns:

Nadin: Az AI-nak nem kell érteni.
Floridi: Az MI-nek nem kell érteni, hogy hatékonyan cselekedhessen.
Esposito: Az MI-nek nem kell érteni ahhoz, hogy társas módon kommunikálhasson.

Ez az egyik elméletileg legeredményesebb párosítás Európa-Európa-Ellen.

Egy második család a **szubjektivitás kiszorításával** foglalkozik. Dukaj az önkéntes szubjektitás elvesztését írja le, mert az algoritmikus delegáció működik. Stiegler a tudásvesztést írja le, mert az automatizálás külsővé teszi a tudást, az érzékelést, a savoir-vivre-t és a savoir-théorisert. Sadin az ítélőképesség elvesztését írja le, mert az MI jogosulttá válik az igazság kimondására. Rouvroy egy még radikálisabb lehetőséget ír le: az algoritmikus kormányzás egyre **inkább elkerülheti az alanyt, mintsem meggyőzze.**

Ez a sorozat így ábrázolható:

Az ész külső kezelése → az ítélkezés → környezeti algoritmikus irányítás → tudásvesztése → az autonóm alany funkcionális szükségletének csökkenése.

Egy harmadik család az **MI-t és a politikai hatalmat** érinti. Bostrom elmélete szerint a szuperintelligencia döntő stratégiai előnyt szerez, és esetleg egyetlen lényt hoz létre. Pokol az egyéni szuperelmét a társadalmilag integrált kollektív szuperintelligencia lehetőségével helyettesíti. Zybertowicz megkérdezi, miért nem képesek a versengő intézmények megakadályozni a fejlődést. Runciman mindkettőt beépíti az államok és vállalatok hosszabb történetébe, mint mesterséges szuperegyéni-ügynököket. Coeckelbergh ezután **mesterséges hatalomként** általánosítja a kérdést.

Egy negyedik család **maga a jövő körül** foglalkozik. Nadin megkülönbözteti az előrejelzést a jóslástól. Esposito megmutatja, hogy a jóslatok képesek előreállítani azokat a jövőket, amiket előre jeleznek. Nowotny azt állítja, hogy ez a performativitás bezárhatja a jövő nyitott horizontját és csökkentheti az emberi időbeli cselekvési lehetőséget. Rouvroy hozzáteszi, hogy az algoritmikus realizmus egy technológiai jövőt nemcsak valószínűnek, hanem **elkerülhetetlennek** is teheti, ezzel gyengítve a politikai képzeletet.

Végül ott van a gazdasági-politikai vonal. Pokol azt kérdezi, hogy a szélsőséges automatizálás alááshatja-e magának a piaci koordináció funkcionális szükségességét. Stiegler megkülönbözteti a foglalkoztatást a lényeges munkától, és egy hozzájáruló utómunka utáni gazdaságot keres. Danaher azt kérdezi, vajon valóban virágzó, munka utáni társadalom épülhet-e. Joler ezzel szemben figyelmeztet, hogy ugyanaz a technológiai folyamat képes rendkívül koncentrált kognitív infrastruktúra tulajdonjogát eredményezni. Ezek nem alternatív válaszok egy kisebb foglalkoztatási kérdésre; versengő elméleteket alkotnak a **munka utáni civilizáció intézményi** formájáról.

Az utolsó rangsorom

Ha egy **nyugati, észak- és dél-európai fejezetet** adnék hozzá a csatolt kötethez, és ugyanolyan szigorúan alkalmaznám az eredetiségi kritériumot, **Luciano Floridit, Elena Espositót, Antoinette Rouvroy-t, Bernard Stieglert és Nick Bostromot** a legvilágosabb első csoportba sorolnám. Mindketten olyan fogalmakat vezettek be, amelyek megváltoztatták a probléma megfogalmazásának fogalmait: *ügynökség intelligencia nélkül; mesterséges kommunikáció; algoritmikus kormányzás; automatikus társadalom/általánosított proletarizáció; és ortogonalitás–instrumentális konvergencia–singleton.*

Közvetlenül mellettük vagy közvetlenül alattuk helyezném **Éric Sadint és Helga Nowotnyt**. Sadin MI-as igazság-tekintélye és tiltó hatalom-elmélete, valamint Nowotny performatív-előrejelzés/nyitott jövő paradoxonája szerintem elég erősek ahhoz, hogy nemzetközileg egyedi elméleti újításoknak, nem pedig pusztán meglévő MI-viták szintézisei.

David Runciman és Mark Coeckelbergh alkotják a következő csoportot: mindketten fontos rendszerszintű átkereteket kínálnak – mesterséges intézményi ügynökséget és mesterséges hatalom/episztemikus ügynökséget –, de eredetiségük inkább a politikai elmélet és az MI radikális újrakombinálásában rejlik, mintsem egy teljesen új általános ontológia megalapításában az MI társadalomban.

Különösen érdekes következtetés, hogy a legmagasabb szintű európai elméletek egyre inkább egy olyan ponthoz koncentrálnak, amely már az Ön kötetében központi jelentőségű: **a döntő kérdés nem csupán az, hogy a gépek intelligensebbek lesznek-e az embereknél.** A mélyebb alternatívák az, hogy az MI átalakítja-e *a társadalmi kogníció* (Pokol), *a szubjektivitás* (Dukaj), *az emberi előrelátó* létezés (Nadin), *az anyagi és episztemikus hatalom* (Joler), *a kommunikáció* (Esposito), *maga az ügynökség* (Floridi), a *kormányzati normativitás* (Rouvroy), *az emberi tudás és az individuáció* architektúráját (Stiegler), vagy a társadalom kapcsolata a *nyitott jövőhöz* (Nowotny). Ez a kibővített összehasonlítás szerintem szilárd elméleti alapot adna egy valóban pán-európai tipológiához, amely az MI társadalom átalakításáról szóló eredeti elméleteket tartalmazza.