

Túl korai lenne: Az AI és a birodalmi logika

KAREN HAO: *Az AI birodalma – Álmok és rémálmok Sam Altman OpenAI-jában*
Budapest: HVG Kiadó, 2025.

Szerzői információ

Dessewffy Tibor, ELTE TáTK Digitális Szociológia Kutatóközpont

Így hivatkozzon erre a cikkre:

Dessewffy Tibor. „Túl korai lenne: Az AI és a birodalmi logika”.

Információs Társadalom XXVI, 1. szám (2026): 94–108.

== <https://dx.doi.org/10.22503/inftars.XXVI.2026.1.14> ==

A folyóiratban közölt művek

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

Nemzetközi Licenc feltételeinek megfelelően használhatók.

I. Bevezetés

„Még korai lenne megmondani”: Az AI és az ítélet felfüggesztése

A gyakran idézett anekdota szerint, amikor a kínai külügyminisztert, Csu En-lajt a hetvenes évek elején a francia forradalom értelméről kérdezték, azt válaszolta: „Még korai lenne megmondani.” Bár ezt sokáig a keleti bölcsesség és a hosszú távú gondolkodás példajaként értelmezték, a kijelentés valójában az 1968-as diáklázadások értékelésére vonatkozott (McGregor 2011).

Ez nem filológiai apróság: a mondat így nem egy lezárt történelmi esemény utólagos megítélésére, hanem egy folyamatban lévő társadalmi átalakulásra utal. Arra a helyzetre, amikor az ítéletalkotás kísértése erős, miközben a strukturális következmények még nem láthatók.

Az AI körüli jelenlegi diskurzus sok tekintetben hasonló helyzetet teremt. A mesterséges intelligenciát övező várakozások és félelmek szélsőséges végpontok által kijelölt horizonton mozognak: az egyik oldalon az emberi munkát felszabadító, háttartalan földi jólétet ígérő technológia, a másikon az egzisztenciális fenyegetésként megjelenő, az emberiségre veszélyt jelentő autonóm gépi értelem víziója. Abban azonban ritkán van vita, hogy az AI, különösen a nagy nyelvi modellek és az ezek köré szerveződő infrastruktúrák, jelentős társadalmi átalakulásokat indítanak el, amelyek hatása nemcsak a gazdaságra, hanem a tudományra, a politikára és a mindennapi élet különböző területeire is kiterjednek.

Karen Hao világsikerű könyve nem semleges elemzés, és ezt a szerző a kötet végén világosan ki is mondja. *Az AI birodalma* egészén végighúzó kritika nem a mesterséges intelligencia technológiáját mint olyat célozza, hanem azt a domináns szilícium-völgyi jövőképet, amelyhez az AI fejlesztése az utóbbi években egyre szorosabban kapcsolódik. Hao számára a tét nem kevesebb, mint annak az elképzelésnek az elutasítása, hogy az AI használata csak egyetlen fejlődési pályán képzelhető el – egy olyan pályán, amely minden más emberi értéket maga alá rendel. Ahogy a könyv epilógusában fogalmaz:

„Az OpenAI-ról és a Szilícium-völgy tágabb víziójáról ebben a könyvben megfogalmazott kritikáim egyáltalán nem a mesterséges intelligencia egészének elutasítását célozzák. Amit elutasítok, az az a veszélyes elképzelés, hogy az AI széles körű társadalmi haszna kizárólag – sőt valaha is – egy olyan technológiai vízióból származhat, amely a magánéletünk, a cselekvőképességünk és az emberi értékünk teljes feladását követeli meg, beleértve munkánk és művészetünk értékét is, egy végső soron birodalmi központosítási projekt szolgálatában” (Hao 2025, 413).

Ez a passzus egyértelművé teszi Hao álláspontját: az „AI-birodalom” nem pusztán leíró metafora, hanem normatív ítélet. A könyv egészének ívét tekintve Hao nem azt igyekszik mérlegelni, hogy vajon mi lesz az AI hosszú távú sorsa, hanem azt, hogy milyen elfogadhatatlan következményregiszterek rajzolódnak ki már most a domináns fejlődési irány mentén. Az OpenAI és a hozzá hasonló szereplők esetében azt állítja: a jelenlegi pálya túl magas társadalmi árat követel, mert az univerzális haszon ideologikus és hamis ígérete valójában egyre inkább az adat, a munka, a tudás és a döntési jogok radikális központosításával jár együtt.

Az én olvasatom ettől némiképpen eltérő. Miközben Hao kritikai lendülete – tekintettel arra, milyen súlyú kérdésekről van szó – érthető és legitim, számomra ezek a folyamatok nyitottabb kimenetelű, nem lezárt kérdések maradnak. Ma még nem látható világosan, mennyiben tekinthető ez a fejlődési irány szükségszerűnek, és pontosan milyen társadalmi átalakulásokat eredményez majd a technológiai dinamika a történeti, intézményi, személyes és politikai döntések bonyolult hálózatában. Nem vitatva az AI korszakos jelentőségét, érdemes felidézni a digitális technológiák korábbi hullámaihoz kapcsolódó irodalom tanulságait is: a rémképek és utópiák, az apokaliptikus víziók és a hurroptimista kinyilatkoztatások viharában a „karaván haladt”, a valóság pedig – bár radikálisan átalakult – eddig nem váltotta be maradéktalanul sem a végletes félelmeket, sem a totális ígéreteket.

Hao alarmista állításaival, azok determinizmusával és véglegességével szemben ezért inkább óvatosabb vagyok. Amit ő határátlépésként jelöl meg, azt én inkább olyan pályaként látom, amelynek kimenetele még nem lezárt – még ha az eddigi tendenciák valóban aggasztóak is. Ez a különbség azonban nem gyengíti Hao könyvét. Éppen ellenkezőleg: kijelöli azt a vitateret, amelyben *Az AI birodalma* nem csupán kritikai állásfoglalásként, hanem komolyan vehető elemzési kiindulópontként is olvasható.

II. Karen Hao – szerzői pozíció

Az AI birodalma megítélése szempontjából nem mellékes kérdés, hogy ki a szerző, és milyen pozícióból beszél. Hao természettudományos alapképzéssel indult, majd rövid ideig egy startupnál dolgozott az egyetem után, mielőtt újságírói pályára lépett volna. Ez a háttér kettős hozzáférést biztosít számára: érti a technológiai nyelvet és a kutatási-fejlesztési logikát, ugyanakkor nem válik egyetlen intézmény vagy vállalat belső narratívájának foglyává.

A könyv empirikus szempontból rendkívül alapos. *Az AI birodalma* több mint háromszáz interjúra épül, körülbelül kétszázhatvan megszólalóval, valamint kiterjedt levelezésekre, belső dokumentumokra és háttéranyagokra. Hao tényfeltáró újságíróként évekig követte az AI-ipar alakulását, különösen az OpenAI történetét, és nemcsak az eseményeket, hanem az iparág kommunikációs működését is testközelből tapasztalta meg. Részletesen érzékelteti, miként zárulnak el bizonyos időszakokban az információs csatornák, majd hogyan nyílnak meg újra krízisek, botrányok vagy hatalmi átrendeződések idején. Ezek a mozgások nem pusztán narratív elemek, hanem az empiria szerves részei: segítenek megérteni, miként működik a hozzáférés, a hallgatás és a kiszivárogtatás politikája a kortárs AI-iparban.

Karen Hao könyvének egyik erénye, hogy következetesen elkerüli a bulváros megközelítést, még akkor is, amikor a tárgyalt valóság erre bőséges muníciót kínál. A hangsúly nem az intimitás kitergetésén, hanem azon van, hogy az egyéni történetek miként ágyazódnak be tágabb kontextusba, és hogyan járulnak hozzá az események megértéséhez. E szabály alól lényegében egyetlen kivétel akad: az a fejezet, amelyben Hao részletesen ismerteti azokat a szexuális zaklatási vádakat,

amelyeket Sam Altman labilis húga fogalmazott meg ellene. Ez a családi tragédiát bemutató rész érzékelhetően kilóg a könyv tónusából, de Hao itt is dokumentumokra, nyilvános nyilatkozatokra és – egymásnak ellentmondó – beszámolókra támaszkodik, és nem próbál végső ítéletet mondani. Arra törekszik, hogy megmutassa: miként válik egy traumatikus történet a hatalom, a hitelesség és az intézményi védelem kérdésévé a zárt és kontrollmániás technológiai elit világában.

Hao plasztikusan rajzolja meg a főbb szereplőket – Sam Altman, Ilya Sutskever, Mira Murati, Greg Brockman és mások alakját – anélkül, hogy egydimenziós hősökké vagy gonosztevőkké redukálná őket. Az egyéni aktorok intézményi helyzetekben mozgó, nagyon is esendő szereplőkként jelennek meg, akiknek döntései és konfliktusai messze túlmutatnak saját személyükön. Hao újságírói erőssége éppen abban áll, hogy képes megmutatni: ezek a döntések többnyire érthetőek az egyéni értékek, félelmek és ambíciók felől, miközben következményeik strukturálisak, tartósak és gyakran visszafordíthatatlanok.

Fontos azonban hangsúlyozni a könyv műfaji korlátait is. *Az AI birodalma* részletes és adat gazdag munka, de alapvetően színvonalas tényfeltáró újságírás, nem pedig szociológiai monográfia. Az olvasó nem talál benne klasszikus társadalomelméleti keretezést, szisztematikus fogalmi apparátust, DOI-számokkal ellátott hivatkozásokot vagy explicit elméleti vitákat. Aki ilyen típusú tudományos feldolgozást vár, óhatatlanul hiányérzettel fogja letenni a könyvet. Hao elsődleges ambíciója nem az elméleti újítás, hanem annak bemutatása, hogyan válik egy technológiai kezdeményezésből politikai-gazdasági jelentőségű megaprojekt – és ebben a műfajban a könyv valóban kimagasló teljesítményt nyújt.

III. Az OpenAI története mint szervezetszociológiai esettanulmány

Az AI birodalma legerősebb és legmeggyőzőbb rétege az OpenAI történetének részletes feldolgozása. Hao elemzése itt nem pusztán egy technológiai vállalat krónikája, hanem egy intézményi átalakulás izgalmas, empirikus rekonstrukciója, amelyben egy altruista küldetésre épülő, idealista startup fokozatosan egy piaci, kapitalista logika által meghatározott megszereplővé válik. A könyv ereje éppen abban áll, hogy ezt az átalakulást nem utólagos moralizálással, hanem konkrét döntések, egyéni szándékok, konfliktusok és strukturális kényszerek mentén mutatja be.

Az OpenAI alapításakor nonprofit szervezatként határozta meg önmagát. A vállalat küldetése az volt, hogy bárki más előtt érje el az általános mesterséges intelligenciát (Artificial General Intelligence, AGI), amelyet az OpenAI „olyan nagymértékben autonóm rendszerekként” definiált, „amelyek az emberi teljesítményt a gazdaságilag legértékesebb tevékenységek többségében felülmúlják”, és amelyek kifejlesztése „az egész emberiség javát” szolgálja. Ez az ethosz – legalábbis kezdetben – nem pusztán marketingfogásként jelenik meg Hao leírásában, hanem a résztvevők által komolyan vett vezérlő csillagként. Ez a küldetés jelentette azt az erkölcsi és technológiai hivatkozási pontot, amely – meggyőződésük szerint – az emberiség számára a legjobb esélyt kínálta arra, hogy az AGI biztonságos módon jöjjön létre.

A szervezet identitása az induláskor részben egy negatív referenciapontra épült: az OpenAI önmagát kifejezetten a nagy technológiai vállalatok, mindenekelőtt a Google adatvezérelt, profitorientált modelljének ellenpontjaként határozta meg. Ez az önmeghatározás azonban már a kezdetektől magában hordozta azokat a feszültségeket, amelyek később meghatározóvá váltak.

Hao részletesen bemutatja, hogyan jelent meg fokozatosan az a dilemma, amely a szervezet egész működését átszövi: miként lehet egyszerre megfelelni egy univerzális, hosszú távú erkölcsi küldetésnek, miközben az AGI-kutatás rendkívüli erőforrásigénye egyre inkább a hatalmas piaci és állami szereplőkhöz köti a szervezetet. A nonprofit forma önmagában nem bizonyult elegendőnek a kutatás finanszírozásához, a számítási kapacitás iránti igény pedig gyorsan olyan nagyságrendet öltött, amely strukturálisan kizárta a hagyományos akadémiai vagy civil működési modelleket.

A könyv egyik fontos erénye, hogy az OpenAI átalakulását nem egyetlen „fordulópontként” ábrázolja, hanem egymásra rakódó döntések sorozataként. A Microsofttal kötött együttműködés, a „capped-profit” modell bevezetése, majd a termékesítés irányába tett lépések nem egyetlen pillanatban rombolták le a kezdeti ideálokat, hanem fokozatosan, komoly és folyamatos belső feszültségekkel alakították át a szervezet működési logikáját. Hao narratívájában ezek a lépések többnyire jól érthetőek és racionálisak az adott helyzetben: a szereplők nem cinikus árulókként, hanem a saját értékeiket és racionalitásukat követő aktorokként jelennek meg.

Az *AI birodalma* egyik figyelemre méltó teljesítménye Karen Hao részéről az, ahogyan az OpenAI vezető alakjait nem ikonokként vagy absztrakt döntéshozókként, hanem különös, gyakran zavarba ejtően esendő figurákként mutatja be. Ezek a portrék nem mellékes anekdoták: Hao világossá teszi, hogy az AI-ipar irányát nem pusztán intézményi kényszerek és piaci logikák határozzák meg, hanem nagyon is konkrét személyiségek, saját félelmeikkel, mániáikkal és világlátásukkal. Hao újságíróként itt van igazán elemében: a személyes részletek nemcsak illusztrálnak, hanem tágabb összefüggéseket magyaráznak.

Sam Altman Hao könyvében az extrém ambíció és a mély egzisztenciális szorongás nyugtalanító elegyeként jelenik meg. Altman nem egyszerűen technológiai optimista, hanem olyan szereplő, aki komolyan számol a civilizáció összeomlásával, miközben aktívan hozzájárul annak felgyorsításához. Hao felvillantja Altman prepper-hajlamait: világvége-szenáriókra való felkészülését, új-zélandi menekülési terveit Peter Thiel társaságában, valamint összekészített „go bag”-et, amelyben pisztoly, antibiotikumok, víz, elemek és sátor lapul – elsősorban egy jövőbeli biológiai katasztrófától tartva.

Ez az apokaliptikus gondolkodás nem válik el Altman technológiai vízióitól. A hosszú élet iránti megszállottság, a 180 millió dolláros befektetés a Retro Biosciencesbe, valamint az érdeklődés a „fiatal vér” kutatása iránt mind ugyanabba az irányba mutatnak: a jövő nemcsak megépítendő, hanem biológiailag is túlélendő projekt. Hao nem kommentálja ezt ironikusan, de az egymás mellé helyezett részletek önmagukért beszélnek.

Altman pszichés sebezhetősége szintén hangsúlyos eleme a portrénak. Hao konkrét epizódokon keresztül mutatja meg a szorongásos tüneteket: a felnőttként is az édesanyját hívó Altman fejfájásait, amelyek mögött agyhártyagyulladás vagy limfómát sejtett; vagy azt az üzleti tárgyalást, ahol pánikroham közeli állapotba került, és félmegtelenül, terpesztett karokkal feküdt a padlón, hogy megnyugodjon. Ezek a jelenetek nem pszichológiai diagnózist szolgálnak, hanem azt az ellentmondást teszik láthatóvá, hogy globális kockázatokat kezelő döntések egy instabil belső világban születnek.

Ehhez társul Altman tudatosan formált fizikai és hedonista önképe. Rendelkezik luxusautó-gyűjteménnyel, benne egy körülbelül 5 millió dolláros Koenigseggel, rendszeresen részt vesz a Burning Man fesztiválon, és Hao beszámolója szerint alkalmi ketamin fogyasztó. Egy munkatársának, annak kirúgását követően egy beszédes epizódban, ketamint és alkoholt ajánlott fel vizsgálatásként. Altman kommunikációs stílusa: a szemtől szembeni konfliktusok kóros kerülése, a folyamatos hazudozás és a manipuláció. A megkérdőjelezhető pénzügyi döntésekhez – amelyekből ő egyénileg profitál – társuló vidám, mindig kisbetűs üzenetek, a projektált nyilvános altruisztikus perszóna tovább erősíti azt a benyomást, hogy a játékos felszín súlyos belső feszültségeket takar.

Hao portréjában Altman figurája egyszerre lenyűgöző és nyugtalanító. Az a világkép rajzolódik ki, amelyben az apokalipszis és a gyors növekedés – startup zsargonban: skálázódás – , a szorongás és a világmegváltó ambíció nem ellentétei egymásnak, hanem ugyanannak a gondolkodásnak a két oldala.

Ha lehet Ilya Sutskever alakja még élesebb kontrasztot hordoz. Hao Sutskevert briliáns, karizmatikus kutatóként mutatja be, aki azonban az AGI-t nem pusztán technológiai kihívásként, hanem szinte transzcendens entitásként érzékeli. A portré tudatosan emeli ki ezeket a különbségeket: az irodában viselt, állatmintás ingeket, a bűjös macskákat, alpakkákat és tűzokádó sárkányokat ábrázoló festményeket, valamint azt a különös szokást, hogy új ötleteket kísérő szellemi feszültséget kézenállásban végzett fekvőtámaszokkal vezeti le.

A könyv egyik legerősebb jelenete a 2022-es vezetői elvonulás, ahol Sutskever egy helyi művésszel készítettett fabábut, amely az „elvetemült, de összehangolt” AGI-t jelképezte. A bábu ünnepélyes elégetése nem ironikus performansz, hanem komolyan vett morális rítus volt: annak deklarálása, hogy az AGI elpusztítása kötelesség, ha veszélyessé válik. A jelenet egyszerre groteszk és hátborzongató, különösen egy olyan szervezet kontextusában, amely közben a világ legerősebb AI-rendszereit fejleszti.

Hao nem rejti véka alá Sutskever paranoiáját sem. A könyv dokumentálja azokat a gondolatkísérleteket, amelyekben Sutskever attól tartott, hogy levágják a kezét, hogy ujjlenyomattal férjenek hozzá az OpenAI titkaihoz, vagy azt javasolta, hogy a vezető tudósokat bunkerszerű védelem alá kellene helyezni. A „Feel the AGI” mantra, amelyet a ChatGPT megjelenése után rendszeresen ismételt, nem marketingeszköz volt, hanem egyfajta belső hitvallás: az AGI jelenlétének emocionális átélése.

Sutskever azon kijelentései, miszerint a mai nagy neurális hálózatok „enyhén tudatosak lehetnek”, tovább erősítik azt a képet, hogy itt nem csupán technikai viták

ról van szó. Hao világosan jelzi: ezek a meggyőződések valós döntési súllyal bírnak a szervezetben belül, különösen a biztonság és a skálázódás dilemmáiban.

Hao portréi azért hatnak ilyen erősen, mert nem redukálhatók pszichológiai érdekességekre. Altman és Sutskever esetében az egyéni jellemhibák és furcsaságok intézményi pályákba csatornázódnak. Az apokaliptikus optimizmus és a misztikus AGI-félelem nem pusztán karakterjegyek, hanem olyan világlátások, habitusok, amelyek strukturálják az OpenAI döntéseit, konfliktusait és végső soron globális hatásait.

Ebben az értelemben Hao könyvének egyik legkellemetlenebb tanulsága az, hogy a mesterséges intelligencia fejlődése nem steril technológiai folyamat. Habár ez nem AI-specifikus, de a jövő alakításában most is nagyon furcsa, gyarló emberek viszik a prímet, és Hao ezt következetesen megmutatja.

Nem véletlen, ha az olvasó ezen a ponton politikai vagy történelmi párhuzamokat vél felfedezni. A könyvön áthúzódó narratív konfliktus – Sam Altman OpenAI-on belüli leváltása a szervezet talán legsikeresebb pillanataiban, majd néhány nappal későbbi visszatérése, az összeesküvők, így Sutskever marginalizálása és eltávolítása – a klasszikus „királydráma” szerkezetét idézi. Az események részletes bemutatása a könyv egyik legerősebb része: az empirikus leírásnál egy akár a bolsevik párt történetéből is ismert szervezeti mintázat rajzolódik ki.

Hao értelmezésében az OpenAI története és a sikertelen „trónfosztás” az eseménytörténeten túl egy általánosabb logikát követ. Azt mutatja meg, hogy egy értékalapú kezdeményezés fokozatosan alkalmazkodik a környezet strukturális kényszereihez, miközben saját eredeti céljai új jelentést kapnak. Ez az átalakulás elkerülhetetlenül mély szervezeti konfliktusokhoz vezet, ahol a korábbi szövetségesek egymással szemben találják magukat – éppen az értékek és ambíciók eltérése miatt. Az átalakulás nem feltétlenül jár együtt az alapértékek teljes feladásával, de azok funkciója megváltozik: az erkölcsi nyelv egyre inkább az intézményi növekedés és túlélés igazolásának eszközévé válik.

A könyv tehát nem csupán az OpenAI történetéről szól, hanem arról is, hogyan működnek a kortárs high-tech szervezetek akkor, amikor egyszerre próbálnak megfelelni etikai elvárásoknak, piaci racionalitásnak és geopolitikai nyomásnak. Hao elemzése itt nem hagy kétséget afelől, hogy egy ilyen helyzetben az idealisztikus értékre épülő stratégiák húzzák a rövidebbet.

IV. AGI mint narratíva, ideológia és szervezőelv

Az OpenAI történetének bemutatása után Hao könyve szükségszerűen túllép egyetlen szervezet keretein, és az AGI köré szerveződő gondolkodásmódot tágabb összefüggésben vizsgálja. Az AGI – az *artificial general intelligence*, vagyis az emberi kognitív képességek széles körét autonóm módon megközelítő mesterséges intelligencia – a kortárs AI-diskurzus egyik központi, ugyanakkor leginkább meghatározatlan fogalma. Potenciális jelentősége ebből fakad: ha megvalósulna, nem pusztán új technológiai eszközt, hanem a gazdasági, politikai és társadalmi rend átalakulásának lehetőségét hordozná. Hao értelmezésében azonban az AGI elsősorban nem

mint konkrét technológiai jövő, hanem mint narratív és ideológiai szervezőelv válik meghatározóvá az AI-iparban.

Az AGI fogalmának homályossága – bizonytalan definíciója, empirikus státusza és folyamatosan eltolódó időhorizontja – nem gyengeségként, hanem erőforrásként jelenik meg ebben a diskurzusban. Éppen ez teszi alkalmassá arra, hogy különböző szereplők eltérő jelentésekkel töltsék meg, és stratégiai célokra használják. Hao implicit állítása szerint az AGI nem elsősorban mint technológiai realitás, hanem mint retorikai és legitimációs eszköz működik: nem az a döntő kérdés, hogy mikor és milyen formában valósul meg, hanem az, hogy mit lehet vele igazolni a jelenben.

Ebben az értelemben az AGI-misszió kettős funkciót lát el. Egyrészt belső motivációs erőként működik: a kutatók és a vezetők számára olyan morális horizontot teremt, amely túlmutat a rövid távú piaci sikereken, és gyakran egzisztenciális jelentőséggel ruházza fel a munkát. Másrészt azonban külső legitimációs eszközzé is válik: alkalmas arra, hogy igazolja a gyors skálázódás kényszerét, az exponenciális fejlődéshez szükséges rendkívüli erőforrás-koncentrációt, valamint azokat a kompromisszumokat, amelyek más kontextusban nehezen lennének védhetők.

Ez a kettősség óhatatlanul felidézi *A Gyűrűk Ura* központi motívumát: a Gyűrű nem pusztán hatalmi eszköz, hanem olyan végcél, amelyhez viszonyulva a szereplők fokozatosan átalakulnak. A Gyűrű ígérete univerzális – rendet, biztonságot, uralmat kínál –, miközben birtoklása és hajszolása lassan feloldja azokat az erkölcsi és közösségi korlátokat, amelyek eredetileg értelmet adtak a küldetésnek. Bár Karen Hao nem él ezzel a párhuzammal, az AGI utáni hajzában hasonló dinamika rajzolódik ki: a cél abszolutizálódása fokozatosan átformálja a szereplők önértelmezését, kockázatérzékelését és döntési horizontját.

Hao bemutatása szerint az AGI köré szerveződő diskurzus az OpenAI-on belül sajátos „végcéllogikát” hoz létre. Ha az AGI elkerülhetetlen, akkor a jelenbeli döntések morális mércéje nem önmagukban, hanem a feltételezett jövőbeli állapothoz való viszonyukban nyer értelmet. Ez a gondolkodásmód könnyen vezet ahhoz, amit a könyv több ponton is az „ends justify the means” logikájaként ír le: a jelenlegi károk, egyenlőtlenségek vagy kizsákmányolási formák relativizálódnak egy feltételezett univerzális jövőbeli haszon fényében. A cél itt nem egyszerűen irányt ad, hanem fokozatosan felmentést is: minél absztraktabbá és távolibbá válik, annál inkább képes elfedni a jelenben meghozott döntések konkrét társadalmi következményeit.

Fontos ugyanakkor, hogy Hao nem vállalkozik az AGI technológiai valószerűségének érdemi megvitatására. Az AGI a könyvben nem jelenik meg mint komolyan mérlegelt jövőbeli technológia; sokkal inkább mint politikai és szervezeti narratíva, amelyet bizonyos szereplők – mindenekelőtt az OpenAI vezetése – tudatosan vagy félig tudatosan mozgósítanak. Ez az elemzési fókusz következetes és a könyv egészéhez jól illeszkedik, ugyanakkor óhatatlanul egyoldalúsághoz vezet: az AGI-t komolyan vevő technológiai optimizmus többnyire naiv vagy manipulatív pozícióként jelenik meg.

Ez a hangsúlyeltolódás különösen akkor válik láthatóvá, amikor az AGI-diskurzuson belüli belső viták kerülnek szóba. Hao érzékenyen írja le a „boomerek” és „doomerek” közötti feszültségeket és ezek szervezeti következményeit, ugyanakkor kevés teret hagy annak a lehetőségnek, hogy az optimista várakozások nem csupán

ideológiai konstrukciókból, hanem technológiai tapasztatokból és fejlesztésekből is táplálkozhatnak.

Ebben az értelemben Hao AGI-értelmezése következetesen szociológiai: nem azt vizsgálja, hogy mi fog megvalósulni, hanem azt, hogy mit tesznek az emberek és intézmények az AGI gondolatával. Ez az erény egyben elő is készíti azt a feszültséget, amely a könyv későbbi fejezeteiben – különösen az „empire” érvelésben – válik igazán meghatározóvá.

Hao elemzésében a skálázódás kényszere és a folyamatosan növekvő számítási kapacitás iránti igény nem valamiféle technológiai „természeti törvényként” jelenik meg, a giga-compute iránti kereslet konkrét szereplők eltérő motivációiból és stratégiai döntéseiből fakad. Az OpenAI esetében a skálázódás logikája kezdetben szorosan összekapcsolódik az AGI-hajszolással, különösen Ilya Sutskever gondolkodásában, ahol a nagyobb modellek nem pusztán hatékonyabb eszközök, hanem az általános intelligencia elérésének előfeltételeiként jelennek meg.

Ez az orientáció azonban fokozatosan találkozik egy másik, legalább ilyen erős hajtóerővel: a piaci verseny logikájával. Sam Altman szerepében Hao azt mutatja meg, hogyan válik a skálázódás részben versenykényszerré, részben üzleti stratégiává. A számítási kapacitás növelése ekkor már nem elsősorban az AGI hosszú távú víziójához, hanem a gyors termékesítéshez, a piaci dominancia megszerzéséhez és a befektetői elvárásokhoz igazodik. A skálázódás így kettős gyökerű folyamattá válik: egyszerre ideológiai és piaci hajtású.

Ez a kettősség kulcsfontosságú Hao érvelésében. A számítási kapacitás iránti exponenciális igény nem szükségszerű technológiai fejlődésként, hanem kontingens döntések sorozataként rajzolódik ki, amelyek idővel strukturális kényszerekké dermednek. Amit kezdetben választásként hoznak meg, az később adottságként jelenik meg a szervezet és az iparág számára. Ebben az értelemben a skálázódás nem csupán technikai kérdés, hanem az AGI-narratíva és a piaci racionalitás találkozási pontja, amely alapvetően meghatározza az AI-ipar további alakulását.

V. Az AI másik oldala – globális egyenlőtlenségek

Az AI birodalma egy másik, fontos rétege azokban a fejezetekben bontakozik ki, ahol Hao szembeállítja az AGI „szent Grálját” hajszoló kaliforniai technológiai elit világát a globális perifériák mindennapi valóságával. A luxusirodák designer bútorai között dolgozó, magukat gyakran morális küldetés hordozóinak tekintő technológiai szereplők világa radikális ellentétben áll azokkal a helyszínekkel, ahol az AI működéséhez szükséges adatmunka ténylegesen zajlik. Ez a kontraszt retorikailag rendkívül erős, ugyanakkor értelmezésre szorul.

Hao empirikus példái – latin-amerikai és afrikai adatfeldolgozók, tartalommoderátorok és -annotátorok történetei – kétségtelenül láthatóvá teszik az AI-ipar globális munkamegosztásának árnyoldalait. Ugyanakkor fontos hangsúlyozni: ezeknek a nyomorúságos élethelyzeteknek az elsődleges oka nem az AI, sőt nem is kizárólag a digitális vállalkozások, hanem a globális kapitalizmus hosszabb történetű, strukturális egyenlőtlenségei. Az AI-ipar ebbe a világba nem „betör”, hanem belép: meglévő

gazdasági és társadalmi viszonyokra épít rá, azokat használja ki, és bizonyos esetekben felerősíti – de nem ő hozza létre őket.

Ez a különbségtétel kiváltképp fontossá válik Hao empátiával átszőtt, újságírói narratívája kapcsán. A venezuelai adatközpont példája, ahol a modelltréninghez kapcsolódó feladatok konkrétan a túlélés feltételét jelentik egy gazdaságilag összeomlott, hiperinflációs környezetben, nem írhatók le egyszerűen kizsákmányolásként. Hasonlóképpen a kenyai leszbikus pár története sem redukálható pusztán a „Big Tech áldozatai” narratívára: ezekben az esetekben az adatközpont ambivalens szerepet tölt be, egyszerre jelent kiszolgáltatottságot és – korlátozott, de reális – cselekvési lehetőséget.

Hao érzékenysége és empátiája kétségtelenül a könyv egyik erőssége. Ugyanakkor felveti azt a kérdést is, hogy ez az empátia nem vezet-e olykor elemzési rövidzárathoz. A globális egyenlőtlenségek drámai eseteinek bemutatása könnyen azt a benyomást keltheti, mintha ezek közvetlenül az AI-ipar „bűnei” lennének, holott sokkal inkább arról van szó, hogy az AI egy már eleve aszimmetrikus világgazdasági térbe ágyazódik be. A morális felháborodás érthető, de társadalomtudományi szempontból fontos megkülönböztetni az okokat, a közvetítőmechanizmusokat és a következményeket.

A könyv eme része így egyszerre rendkívül erős és problematikus. Erős, mert konkrét emberi sorsokon keresztül teszi láthatóvá az AI-ipar globális működését, és problematikus, mert az empirikus esetek súlya olykor elhomályosítja a strukturális elemzést. Az adatközpont, a láthatatlan munka és a környezeti terhelés valóban szerves részei az AI-iparnak, de ezek értelmezése csak akkor válik analitikusan termékennyé, ha világosan látjuk a szereplők számára felmerülő alternatívákat, elválasztjuk egymástól a globális kapitalizmus általános logikáit és az AI-ipar ezekhez való specifikus hozzájárulását.

Az *AI birodalma* a környezetkárosítás dimenzióját Hao nem elvont fenntarthatósági diskurzuson keresztül bontja ki, hanem kifejezetten drámai lépték-összehasonlításokkal dolgozik. A nagyméretű AI-modellekhez szükséges adatközpontok energiaigényét és vízfogyasztását olyan analógiák segítségével teszi érzékelhetővé, amelyek egyértelműen túlmutatnak a „digitális” iparág intuitív képzetén. A könyvben visszatérően megjelenik az a gondolat, hogy egyetlen hiperskálázott adatközpont energiaigénye városi léptékű infrastruktúrákkal mérhető össze, nem pedig kutatólaborokéval.

Hao több helyen idéz olyan becsléseket, amelyek szerint az AI-megakampuszok energiaigénye elérheti az 1000–2000 megawattos nagyságrendet, vagyis egyes nagyvárosok éves áramfogyasztásával összemérhető terhelést jelenthet. Hasonlóképpen a Microsoft–OpenAI-együttműködésben tervezett Stargate szuperszámítógép kapcsán a könyv olyan projekciókat említ, amelyek New York City átlagos áramigényével vetik össze egyetlen létesítmény potenciális fogyasztását. Ezek az összehasonlítások nem technikai pontosságra, hanem érzékeltetésre törekednek: arra, hogy világossá tegyék, az AI-fejlesztés ipari, nem pedig marginális energetikai kérdés.

A vízfelhasználás kapcsán Hao még konkrétabb példákat hoz. Az adatközpontok hűtéséhez szükséges ivóvíz-felhasználás különösen konfliktusossá válik olyan régiókban, ahol a vízhiány már eleve társadalmi feszültségeket okoz. Chile esetében a könyv részletesen bemutatja a Santiago környékére tervezett adatközpont ügyét,

ahol egy aszály sújtotta térségben másodpercenként száz literes nagyságrendű ivóvíz-felhasználás váltotta ki a helyi közösségek ellenállását, végül a beruházás elutasítását. Hasonló konfliktus rajzolódik ki Uruguay példáján keresztül is, ahol egy súlyos ivóvízválság idején tervezett adatközpont napi több millió gallon vízfelhasználása vált politikai botránnyá, és kényszerítette ki a projekt áttervezését.

Ezek a példák Hao narratívájában nem elszigetelt botránnyokként jelennek meg, hanem annak illusztrációiként, hogy az AI-ipar konkrét helyeken, konkrét közösségekre terhel rá globális infrastruktúra-igényeket. A hangsúly nem azon van, hogy egy-egy szám „pontosan” mennyi, hanem azon, hogy az erőforrás-felhasználás versenyhelyzetbe kerül más alapvető társadalmi szükségletekkel: ivóvízzel, lakossági energiaellátással, mezőgazdasági felhasználással.

Hao környezeti érvelésének ereje éppen ebben a léptékváltásban rejlik. Az AI itt végleg elveszíti „súlytalan” digitális jellegét, és olyan iparágként jelenik meg, amely energia-, víz- és földhasználati döntéseken keresztül avatkozik be társadalmi folyamatokba. Ugyanakkor – összhangban a könyv egészének logikájával – ezek a költségek sem önmagukban, hanem a globális kapitalizmus meglévő strukturális egyenlőtlenségeibe ágyazva válnak igazán értelmezhetővé. Az AI nem hozza létre ezeket a konfliktusokat, de új intenzitást és új léptéket ad nekik.

VI. Az Empire mint elméleti keret

Karen Hao könyve négy, egymásra épülő rétegen keresztül halad: az OpenAI szervezeti történetének részletes feltárásától az AGI hajszolása köré szerveződő ideológiai–technológiai komplexum elemzésén át a globális társadalmi és környezeti következmények bemutatásáig, majd végül ezek szintéziseként jut el az „AI-birodalom” fogalmához. Ez az utolsó réteg minőségi váltást jelent: Hao itt már nem elsősorban empirikus eseteket sorol, hanem egy átfogó metaforán keresztül próbálja megragadni az AI-ipar működésének logikáját.

Az „empire” fogalma ebben az értelemben nem pusztán retorikai dísz, hanem értelmezési kísérlet. Hao azt állítja, hogy az OpenAI és a hozzá hasonló szereplők működésében olyan mintázatok jelennek meg, amelyek történetileg a birodalmak működésére emlékeztetnek. Nem klasszikus értelemben vett uralomról vagy közvetlen politikai kontrollról van szó, hanem egy olyan hatalmi konfigurációról, amely gazdasági, technológiai és kulturális eszközökön keresztül fejt ki hatását. A könyv négy ilyen dimenziót azonosít, amelyek együtt adják ki az „AI-birodalom” logikáját.

Az első dimenzió az ideológiai igazolás. Hao szerint az AGI univerzális, homályosan megfogalmazott küldetése – „az emberiség javát szolgáló általános mesterséges intelligencia” – funkcionálisan hasonló szerepet tölt be, mint a történelmi birodalmak „civilizációs missziói”. Ez a narratíva egyszerre mobilizál tehetséget és tőkét, valamint morális keretet ad olyan döntéseknek, amelyek rövid távon konkrét társadalmi károkkal járnak. Az „jó birodalom vs. rossz birodalom” logika – például a Kínával szembeni verseny hangsúlyozása – nem mellékes retorikai elem, hanem az intézményi cselekvés egyik legitimációs forrása.

A második dimenzió az erőforrások kisajátítása. Hao ide sorolja nemcsak az internetes adatok tömeges felhasználását, hanem a természeti erőforrásokat is: az energia- és vízigényt, valamint az infrastruktúrák térbeli koncentrációját. Az AI-ipar ebben az olvasatban olyan erőforrásokra épít, amelyek felett nem rendelkezik közvetlen tulajdonjoggal, mégis globális léptékben vonja őket működésébe. Ez a folyamat nem feltétlenül jogellenes, de strukturálisan hasonló ahhoz, ahogyan a történelmi birodalmak perifériáikról vonták be az erőforrásokat a központ működtetéséhez.

A harmadik dimenzió a munka strukturális kihasználása. Hao itt egyszerre beszél a globális adat- és tartalommoderációs munkáról, valamint az automatizáció hosszabb távú munkaerőpiaci hatásairól. A „ghost work” nem elszigetelt anomália, hanem az AI-fejlesztés egyik alapfeltétele, miközben a létrejövő rendszerek éppen a munkaerő kiváltásának ígéretére épülnek. Ebben a kettősségben Hao olyan hatalmi aszimmetriát lát, amely túlmutat az egyedi visszaéléseken, és a globális munkamegosztás újraformálását jelzi.

A negyedik dimenzió a tudás és értelmezési keretek monopolizálása. Hao szerint az AI-jal kapcsolatos szakértelem és kutatási kapacitás fokozódó koncentrációja nemcsak gazdasági, hanem kulturális hatalmat is jelent. Amikor a legfontosabb kutatók és narratívák vállalati keretek közé szorulnak, a technológiáról alkotott közbeszéd szükségszerűen ezen intézmények szűrőjén keresztül formálódik. Ez a tudásmonopólium nem formális cenzúra révén működik, hanem a láthatóság, a hozzáférés és az autoritás egyenlőtlen elosztásán keresztül.

Hao birodalommetaforája ebben az értelemben nem történelmi egyenlőségjelet tesz múlt és jelen közé, hanem szerkezeti hasonlóságokra mutat rá. A könyv ugyanakkor nem bontja ki részletesen az analógiát: az egyetlen konkrét történelmi párhuzam a Brit Kelet-indiai Társaság, és az összevetés is inkább vázlatos marad. Ez a megközelítés ugyan lehetővé teszi, hogy az „empire” fogalma provokatív értelmezési keretként működjön, de nyitva hagyja a kérdést, mennyiben beszélhetünk valóban új típusú birodalomról, és mennyiben a digitális kapitalizmus ismert logikáinak újracsomagolásáról.

A metafora akkor válik igazán élessé – és sebezhetővé –, amikor az általa javasolt megoldásokkal vetjük össze. Ezen a ponton derül ki, hogy az „AI-birodalom” mennyiben diagnózis, és mennyiben normatív túlzás.

Az *AI birodalma* zárófejezetében (Epilógus: *Hogyan omlik össze a birodalom*) Karen Hao nem elégszik meg a diagnózissal, hanem normatív irányba lép tovább. Megoldási javaslatait az „AI-birodalom” alapját képező hatalmi koncentráció lebontása köré szervezi. Kiindulópontja szerint a hatalom az AI-iparban három fő tengely mentén sűrűsödik: a tudás, az erőforrások és a befolyás terén. A birodalom lebontása ezért csak e három dimenzió együttes újraelosztásával képzelhető el.

Hao szerint az AI-birodalom egyik legfontosabb pillére a tudás termelésének és értelmezésének vállalati koncentrációja. Ennek ellensúlyozására a független kutatás megerősítését javasolja. Ide tartozik a nem vállalati finanszírozású AI-kutatás támogatása, különösen olyan intézmények esetében, amelyek nem a nagyméretű nyelvi modellek skálázásában látják a fejlődés egyetlen útját, hanem alternatív, adat- és energiahatékonyabb megközelítéseket keresnek.

Hao emellett hangsúlyozza a közösségi alapú, feladat specifikus AI-alkalmazások jelentőségét, amelyek nem univerzális technológiai megoldásokat ígérnek, hanem konkrét közösségek problémáira reagálnak. Ehhez szorosan kapcsolódik a vállalati modellek független értékelésének igénye: a társadalom ne kizárólag a cégek saját állításaira támaszkodjon az AI-rendszerek képességeiről és kockázatairól. A tudás újraelosztásának feltétele tehát az átláthatóság – a betanítási adatok, a compute-infrastruktúra és a modellek valós hatásainak nyilvánossága.

A második nyaláb az AI-ipar anyagi alapjait célozza. Hao szerint az AI-birodalmak hatalmukat az erőforrások – adat, munka, energia, tőke – felhalmozásával biztosítják, ezért ezek újraelosztása nélkül minden más reform felszínes marad. Ennek egyik kulcseleme a munkaügyi védelem megerősítése: nemcsak az adatannotátorok és tartalommoderátorok esetében, hanem azoknál az alkotóknál is, akiknek jogvédett munkái betanítási adatokká váltak. A szakszervezetek és kollektív érdekvépviseletek szerepe Hao szerint kulcsfontosságú az automatizációval szembeni alkupozíció megteremtésében.

Ehhez kapcsolódik az IP-szabályozás frissítése és az ellátási láncok átláthatóvá tétele. Hao mellett érvel, hogy a betanítási adatok forrásainak nyilvánossága és az infrastruktúra-beruházások láthatóvá tétele megnehezítené az extraktív, „észrevétlen” erőforrás-kisajátítást. Végül az AI-termékek fogyasztóvédelmi jellegű felügyeletét is szükségesnek tartja, amely nem elméleti kockázatokat, hanem a rendszerek tényleges társadalmi hatásait vizsgálná.

A harmadik tengely a befolyás kérdésére irányul. Hao szerint az AI-birodalom fennmaradását nemcsak az erőforrások koncentrációja, hanem a technológia köré épülő „mágikus” aura is biztosítja. Ennek lebontásához széles körű oktatásra van szükség: nem kódoló-képzésre, hanem annak megértésére, hogyan működnek az AI-rendszerek, mik a korlátaik, és milyen világnézetek, feltételezések állnak a fejlesztőik mögött. A cél az, hogy az AI elveszítse megkérdőjelezhetetlen autoritását, és társadalmi viták tárgyává váljon.

Ebben az értelemben Hao számára az AI governance alapvető kérdése nem technikai, hanem politikai: a mesterséges intelligencia koncentrálja vagy újra elosztja-e a hatalmat. A birodalmi ideológia kívánatos lebontása pedig csak akkor lehetséges, ha ez a kérdés expliciten megjelenik a szabályozásban, az oktatásban és a közbeszédben.

Normatív erő, strukturális kihívások

Hao megoldási javaslati koherensek és kétségkívül kívánatosak lennének. Ugyanakkor már első olvasatra is felmerül a kérdés, mennyiben tekinthetők ezek reálisan megvalósítható programnak, és mennyiben inkább egy kritikai horizont kijelölésének. A tudás, az erőforrások és a befolyás újraelosztása nem pusztán szakpolitikai döntések kérdése, hanem mélyen beágyazott gazdasági érdekeket, szabályozási korlátokat és – nem utolsósorban – társadalmi közönyt érint.

Karen Hao megoldási javaslati normatív értelemben nehezen vitathatók. A tudás, az erőforrások és a befolyás újraelosztására épülő program belsőleg követke-

zetes, és reagál azokra a problémákra, amelyeket *Az AI birodalma* korábbi fejezetei részletesen feltárnak. Ha az AI valóban birodalmi logika szerint működik, akkor a hatalom koncentrációjának megtörése nem opcionális, hanem szükségszerű célkitűzés.

A könyv eme pontján azonban feszültség keletkezik a diagnózis és a terápia realitása között. Hao javaslatai ugyanis implicit módon feltételezik, hogy a hatalom újraelosztása elsősorban szabályozási és intézményi akarat kérdése. A probléma az, hogy az általa leírt birodalmi logika éppen azt mutatja meg, miért nem működik ez az előfeltevés magától értetődően.

A tudás újraelosztására irányuló javaslatok például egy olyan tudományos és finanszírozási környezetben ütköznek korlátokba, ahol a legnagyobb kutatási infrastruktúrák, adatvagyonok és számítási kapacitások már eleve vállalati kézben vannak. A független kutatás támogatása normatívan kívánatos, de strukturálisan marginális marad mindaddig, amíg nem képes versenyezni a skálázódás logikájával. A kérdés itt nem az, hogy létezhet-e alternatív AI-kutatás, hanem az, hogy milyen feltételek mellett válhat releváns ellenpólussá, és erre Hao kevésbé ad választ.

Hasonló feszültség figyelhető meg az erőforrások újraelosztásának programjában. A munkaügyi védelem, az IP-szabályozás megerősítése és az ellátási láncok átláthatósága mind olyan célok, amelyek évtizedek óta napirenden vannak a globális kapitalizmus más ágazataiban is – mérsékelt sikerrel. Az AI-ipar sajátossága éppen az, hogy gyors technológiai ciklusai és geopolitikai beágyazottsága miatt különösen ellenáll az ilyen típusú beavatkozásokkal szemben. A „kemény érdekek” itt nem absztrakt fogalomként, hanem állami versenypolitikák, nemzetbiztonsági narratívák és befektetői elvárások formájában jelennek meg.

A befolyás újraelosztására épülő oktatási és demisztifikációs program talán a legszimpatikusabb, de egyben a leginkább alultervezett elem. Hao joggal hangsúlyozza, hogy az AI körüli „mágia” fenntartja a technológia társadalmi autoritását. Ugyanakkor kevésbé reflektál arra, hogy a széles körű technológiai oktatás és kritikai tudás terjesztése milyen mértékben ütközik társadalmi érdektelenségbe. A legtöbb felhasználó számára az AI nem politikai kérdés, hanem kényelmi eszköz, és ez a közöny önmagában is stabilizáló tényezővé válik.

Ezen a ponton válik igazán láthatóvá a könyv egyik alapfeszültsége. Hao rendkívül meggyőzően mutatja meg, hogy az AI-ipar működése strukturális tényezők által vezérelt rendszer. Megoldási javaslatai viszont gyakran úgy szólnak, mintha ezek a kényszerek részben megkerülhetők lennének normatív beavatkozásokkal. Ez nem teszi irrelevánssá a javaslatokat, de eltolja őket egy másik regiszterbe: program helyett inkább morális iránytűként értelmezhetők.

Ebben az értelemben *Az AI birodalma* nem annyira cselekvési tervet, mint inkább kritikai horizontot kínál. Hao könyvének legnagyobb erénye abban áll, hogy világossá teszi: a mesterséges intelligencia jövője nem technológiai szükségszerűség, hanem politikai és társadalmi döntések, konfliktusok és harc eredménye. Hogy ezek a döntések milyen irányba mozdulnak el, bár erre vonatkozó intuíciónk lehetnek, azt ma még valóban „túl korai lenne megmondani”.

Összességében *Az AI birodalma* ritka módon képes egyszerre több olvasóközöniséget megszólítani. A mesterséges intelligenciával most ismerkedők számára átfogó,

olvasmányos belépést kínál az iparág társadalmi és politikai kérdéseibe, míg a témával mélyebben foglalkozó kutatók számára is izgalmas olvasmány marad empirikus gazdagsága és provokatív felvetései miatt. Hao könyvével nem szükséges minden ponton egyetérteni – sőt bizonyos állításai kifejezetten vitára hívnak –, de éppen alapos ismeretanyaga és következetesen végiggondolt, gyakran szándékoltan éles felvetései teszik olyan munkává, amellyel érdemes komolyan foglalkozni.

Irodalom

McGregor, Richard: What Zhou meant by ‘Too early to say’, *The Financial Times* 2011. június 11.
Hao, Karen. *Empire of AI: Dreams and nightmares in Sam Altman’s OpenAI*. Penguin Group, 2025.