

Pokol Béla

AI-jövő társadalmai

(a ChatGPT5.2 deep search segítségével)

Tartalom

Kurzweil.....1
Harari..... 25
Crawford.....44
Lee és Qiufan..... 56

Kérjük, részletesen írd le Ray Kurzweil 2024-es új könyvét, a "The Singularity is Nearer"-t.

Ray Kurzweil „The Singularity Is Nearer” (2024) művének áttekintése

Ray Kurzweil 2024-es folytatása, a Szingularitás közelebb: Amikor az AI-val egyesülünk, újra és frissíti 2005-ös könyvének, a The Singularity Is Near-nak víziós témáit. Azt állítja, hogy az exponenciális technológiai fejlődés – különösen a mesterséges intelligencia (MI), a biotechnológia és a nanotechnológia területén – egy jövőbeli "Szingularitáshoz" vezet, ahol az emberi intelligencia egyesül a gépi intelligenciával. Kurzweil előrejelzi, hogy a 2040-es évek közepére az emberek jelentősen bővítik kognitív képességeiket azáltal, hogy agyukat összekapcsolják az MI-vel, áttöréseket érnek el a mai képzeleten felülmúló hosszú életben, tudásban és képességekben.

Kurzweil új könyve egyszerre útterv és előrehaladási jelentés. Vizsgálja, mennyit jutottunk el 2005 óta, megduplázza a kulcsfontosságú előrejelzéseket (például az emberi szintű MI 2029-re és teljes ember-AI fúziót 2045-re), és foglalkozik a kortárs lehetőségekkel és szorongásokkal ezekkel a technológiákkal kapcsolatban. Az alábbiakban részletes, fejezetről fejezetre szóló

összefoglaló található a *The Singularity Is Nearer-ről*, majd a főbb előrejelzések, technológiai előrejelzések, fejlődő ötletek, valamint Kurzweil által tárgyalt filozófiai és társadalmi vonatkozásai áttekintése. Bemutatjuk azt is, hogyan reagál Kurzweil a kritikusokra, és hogyan fogadták a könyvet a tech közösségben és a médiában.

Fejezetről fejezetre összefoglaló

1. fejezet: Hol vagyunk a hat szakaszban?

Kurzweil azzal kezdi, hogy áttekinti az intelligencia evolúciójának hat korszakát, amelyeket először korábbi műveiben vázolt fel. Ezek az időszakok a fizika és kémia (1. korszak) és az élet megjelenése (2. korszak) irányát követik nyomon az agy megjelenésén (3. korszak) és a technológián (4. korszak), majd az ember–mesterséges intelligencia integrációjának közelgő korszakaiig (5. korszak), majd végül az intelligencia univerzumon belüli elterjedéséig (6. korszak). 2024-ben a 4. korszak végéhez közelítünk, az információs technológia felgyorsítja az emberi fejlődést, és az 5. korszak (agy-számítógép interfészek) csak most kezd kialakulni. Kurzweil kiemeli a jelenlegi MI eredményeit – például *a Go* és *a Jeopardy!* az orvostudomány fejlődéséig – mint annak a jele, hogy a következő szakasz küszöbén vagyunk. Megjegyzi, hogy az MI rendszerek gyorsan fejlődnek, és előrejelzi, hogy *2029-re egy MI átmegegy egy robusztus Turing-teszten*, ami azt jelenti, hogy megkülönböztethetetlenül beszélhet az embertől. Ez a fejezet megteremti a színpadot, a történelmet exponenciális intelligencianövekedésként ábrázolva, és felkészíti az olvasót arra a "elképesztő kognitív fejlődésre", amelyet Kurzweil szerint a Szingularitás hozni fog.

2. fejezet: Az intelligencia újjáépítése

Itt Kurzweil magának a mesterséges intelligencia fejlődésének bemutatására merül – hogyan *"újjátaláljuk meg azt az intelligenciát, amelyet a természet adott nekünk, egy erősebb, digitális alapon keresztül"*. Történelmi körútja az MI-kutatásoknak, bemutatva az elmúlt fél évszázad versengő megközelítéseit, és hogy miért volt lassú a korai előrehaladás. Kulcsfontosságú érv az, hogy az AI legutóbbi áttöréseit (különösen a mélytanulási neurális hálózatokban) a számítási teljesítmény exponenciális növekedése tette lehetővé. Sok évtizedekkel ezelőtti MI-ötlet nem a betekintés hiánya miatt, hanem a megfelelő számítási képesség hiánya miatt – *"előrelátó ötletek, de nem működőek, amíg jobb hardver nem érkezett"* addig. Most, hogy a számítástechnikai erőforrások robbanásszerűen megnöttek, az MI végre emberszerű mintafelismerést és gondolkodást érhet el analógiával, akárcsak az emberi agy. Kurzweil élénk példákat használ, például Charles Darwin analógiáit (a biológiai evolúciót a geológiai folyamatokkal való összehasonlítás) arra, hogy szemléltessene az analógia intelligencia erejét. A fejezet nagyobb témája, hogy biológiai "állatokból" *"transzcendens lényekké"* váltunk, akiknek gondolatai és identitásai már nem a biológiánk által bilincsben vannak. A 2020-as években az emberiség ennek az átalakulásnak a végső szakaszába lép: szó szerint egy szintetikus elmét alkotunk, és hamarosan összeolvadunk vele.

3. fejezet: Ki vagyok én?

Ez a filozófiailag gazdag fejezet a tudatosságot és az identitást vizsgálja egy olyan korszakban, amikor a gépek gondolkodhatnak, és az emberek bővíthetik az elméjüket. Kurzweil megkérdőjelezi azt a nézetet, hogy a tudatosság kizárólag emberi tulajdonság. Rámutat a kutatásokra, amelyek szerint sok állat rendelkezik öntudattal és szubjektív tapasztalattal

(például *hivatkozva a Cambridge-i Tudatossági Nyilatkozatra* az állati tudatosságról). Az állati elme státuszának emelésével precedenst teremt a gépi elmék gondolkodásához. Kurzweil azt állítja, hogy ha egy nem biológiai intelligencia ugyanazt a bonyolultságot és viselkedést mutatja, mint egy tudatos ember vagy állat, *akkor "szubjektív belső életet is adna neki"* – vagyis egy elég fejlett MI *valóban* tudatossággal rendelkezik, nem csak szimulálná azt. Ez elgondolkodtató kérdésekhez vezet: Hogyan ismerjük el vagy adjuk meg a jogokat az intelligens gépeknek? Milyen etikai keretekre van szükség, amikor egy mesterséges intelligencia vagy feltöltött emberi elme azt mondja: "Érzem"? Kurzweil előrevetíti az MI személyiségéről szóló vitákat, és hangsúlyozza a *"felelősségünket... hogy kialakítsuk az etikai szellemiséget más nem-emberi tudatosság formái számára"*, amelyeket mi létrehozunk. Elmélkedik az élet és az elme csodálatos valószínűtlenségén is az univerzumban. A Nagy Bumm utáni apró fizikai állandó változások megakadályozták volna csillagok, bolygók vagy élet kialakulását; egy csillagász számítása szerint a Föld életet hordozó összetettségének véletlenszerű kialakulásának esélye összehasonlítható egy tornádóval, amely egy Boeing 747-est szerel össze egy roncstelepből. Ez a kozmikus csodálat kvázi-spirituális dimenziót ad a fejezetnek. Végző soron a "Ki vagyok én?" azzal küzd, mi határozza meg az éntünket, miközben közeledünk az elmefeltöltés és ember-mesterséges intelligencia hibridek korához. Kurzweil azt javasolja, hogy a tudatosság egy *folytonosság*, nem egy bekapcsolva/kikapcsolódó tulajdonság, amely egy adott alaphoz kötött, és hogy az intelligencia bővítése megváltoztathatja, hogyan határozzuk meg identitásunkat.

4. fejezet: Az élet exponenciálisan jobb lesz

Kurzweil egy vidám, adatvezérelt emberi fejlődés szemléletére tér át. Ebben a fejezetben olyan gondolkodókkal áll, mint Steven Pinker (*Enlightenment Now*) és Peter Diamandis (*Bőség*), és azt állítják, hogy a komor népszerű megítélések ellenére szinte minden objektív emberi jóléti mutató drámaian javult évtizedek és évszázadok alatt. Katalogizálja a "végtelen fejlődési mátrixokat": növekvő globális írástudás, egészség és vagyon; a szélsőséges szegénység, a gyermekhalandóság, az erőszak és a munkaidő jelentősen csökkenő aránya; fokozott hozzáférés az oktatáshoz és az információhoz; tisztább levegő és víz; és így tovább. Kurzweil szerint ezek közül a pozitív trendek közül sok a technológia miatt gyorsul. Itt bemutatja védjegyét, a Gyorsító Hozamok Törvényét (LOAR) – az információalapú technológiák visszacsatolással vezérelt exponenciális fejlesztését. A technológiai fejlődés nem egyenes vonalban halad, hanem egyre meredekebb görbében, mert minden fejlődési generáció eszközöket ad a következő fejlesztések megkönnyítéséhez. Kurzweil például a számítási költségek drámai csökkenését (egy modern laptop milliószor erősebb, mint egy 1970-es évekbeli számítógép, miközben sokkal olcsóbb) és a mezőgazdaságban bekövetkezett áttörések (például MI-optimalizált függőleges gazdaságok, amelyek kevesebb földdel és vízzel több élelmiszert termelnek). Még a múltbeli technológiai változásoknak is tulajdonítja a széleskörű társadalmi változásokat, például azt, hogy a 20. századi háztartási gépek milliókat szabadítottak meg a fáradságtól, és lehetővé tették számukra, hogy hatalmas számban lépjenek be a munkaerőpiacra. A cím állítása, miszerint *"az élet exponenciálisan jobb"* lesz, statisztikák és grafikonok is támasztják alá, néha évszázadokat vagy évezredekkel ölelnek fel. (A könyv egy kis humorral megjegyzi, hogy a bűncselekmény ma alacsonyabb, mint mondjuk a 14. században – bár az emberek nem találják ezt a párhuzamot megnyugtatónak.) Kurzweil elismeri, hogy problémák továbbra is fennállnak, de kijelenti, hogy a technológia összetett előnyei továbbra is megoldják az alapvető emberi kihívásokat. Például azt képzelem el, hogy a laboratóriumban termesztett hús helyettesíti a gyári gazdálkodást a környezeti károk csökkentése érdekében (évente több mint 70 milliárd állatot vágnak le élelmiszernak, ami várhatóan csökkenni fog), és a 3D nyomtatást, amely forradalmasítja a gyártást (még az átültetésre alkalmas szerveket is hamarosan 3D-nyomtatásra

lehet nyomtatni). Pinker és Diamandis optimizmusát saját LOAR keretrendszerével párosítva Kurzweil azt állítja, hogy a következő két évtizedes exponenciális növekedés még mélyebb javulást hoz – potenciálisan *felszámolva a szegénységet és a betegségeket*, valamint tovább terjesztve a demokráciát és az emberi jogokat világszerte.

5. fejezet: A munkahelyek jövője – jó vagy rossz?

Ez a fejezet az egyik legközvetlenebb szorongást foglalkoztatja az MI kapcsán: elveszi a munkánkat? Kurzweil válasza árnyalt. Először elismeri a zavart. Sok munkahelyet *megszüntet* az MI és az automatizálás, és hamarosan. Például megjegyzi, hogy az Egyesült Államokban körülbelül 4,6 millió ember (a munkaerő 2,7%-a) dolgozik vezetőként, és *"szinte biztos, hogy sokan közülük elveszítik állásukat, mielőtt egyébként nyugdíjba vonultak volna"* az önvezető járművek miatt. Hasonlóképpen, az MI által vezérelt rendszerek egyre inkább uralják majd a gyártást a 2030-as évekre. Kurzweil történelmi párhuzamokat követ el: ahogy a mezőgazdaság 1900-ban a legtöbb munkást ma már kevesebb mint 2%-ra vált, és a gyártás 1970-ben az amerikai munkahelyek 25%-áról ma már kevesebb mint 8%-ra csökkent, az MI is új szektorokba fogja áthelyezni a munkaerőt, amelyeket alig tudunk elképzelni. Lényeges, hogy *új munkahelyek és teljesen új iparágak jelennek meg*, ahogy az minden korábbi technológiai forradalom során is történt. Például az internet felemelkedése olyan munkahelyeket teremtett, mint az "alkalmazásfejlesztő" és a "közösségi média menedzser", amelyek egy generációval ezelőtt ismeretlenek voltak. Kurzweil hangsúlyozza, hogy a termelékenység növekedése végső soron gazdagságot emel, és keresletet teremt új munkatípusok iránt, még ha az átmenet rázós is. Statisztikákra hivatkozik, hogy ma *világszerte több munkahelyünk van, mint valaha*, és jóval magasabb az egy főre jutó jövedelem, mint 100 évvel ezelőtt az inflációval korrigálva. Az átmenet irányítása érdekében Kurzweil merész társadalmi innovációkat szorgalmaz, mint például az egyetemes alapjövedelem (UBI). Előre jelezi, hogy valamilyen UBI *"a 2030-as években kezdődik"* a kitelepített munkavállalók enyhítésére, és bár a korai UBI programok *"akkor nem lesznek megfelelőek"*, idővel bővülni fognak az MI által teremtett hatalmas gazdagsággal együtt. Azt is javasolja, hogy újra kell definiálnunk, hogyan mérjük a gazdasági fejlődést, mivel sok digitális áru és szolgáltatás (vagy az MI által elért termelékenységi javul) nem van jól rögzítve a hagyományos mutatókban. Végül Kurzweil optimista oldalon áll: ahogy a 19. századi ludditák végül nem erőszakkal, hanem a technológia által teremtett nagyobb jólét által *"elpusztították"* a társadalom számára, úgy véli, hogy az MI "közelgő bősége" hasonlóan legyőzi a mai félelmeket. Ennek ellenére előrelátásra és *"nagyon szándékos"* erőfeszítésekre szorgalmazza a munkavállalók újraképzését és a rövid távú zavarok mérséklését politikai úton keresztül – elismerve, hogy *a Szingularitás felé vezető úton óvatosan kell járni*.

6. fejezet: A következő 30 év az egészségben és jóllétben

Kurzweil a biomedicinához és az emberi élettartamhoz fordul, előrejelzést adva, hogyan alakítják át a következő három évtizedben az egészséget a technológiák. Megfigyeli, hogy az orvoslás ma *továbbra is "kusz" és pontatlan* – az orvosok gyakran próbálkozás-hibás és népszerűlátásokat alkalmaznak, nem pedig igazán egyéni tudományt. Azonban az orvostudomány gyorsan információs technológiává válik, amely képes ugyanolyan exponenciális nyereséget hozni, mint a számítástechnika. Az olyan fejlesztések, mint az MI-alapú gyógyszerfelfedezés és a genetikai szerkesztés, már most is feltűnő eredményeket hoznak. A könyv példákat hoz, például egy 2019-es ausztrál projektet, amely MI-t használt egy "turbófeltöltős" influenzavakcina tervezésére a szokásos idő töredékében, valamint egy 2020-as MIT MI rendszer, amely 107 millió molekulát szűrt egy erős új antibiotikumot azonosítani (ami napok alatt elérte azt, amit az embernek évekbe telt volna). Kurzweil kiemeli, hogy a

COVID-19 járvány idején az MI mindössze 63 nap alatt segített egy vakcina kidolgozásában, míg a hagyományos vakcina fejlesztése általában 5–10 évig tartott. Azt jósolja, hogy a 2020-as évek végére szinte minden orvosi diagnosztikát MI végzi majd – például az MI-k már szakértői pontossággal vizsgálják a röntgensugarakat és orvosi képeket. A kezelés is sokkal pontosabb lesz, ahogy megtanuljuk újraprogramozni a biológiát. Kurzweil egyik központi gondolata itt a "hosszú távú menekülési sebesség" – egy kifejezés Aubrey de Grey gerontológustól –, amely arra a fordulóra utal, amikor az orvostudomány évente több mint egy évvel meghosszabbítja az emberi élettartamot, így a fennmaradó élettartamunk gyakorlatilag nem csökken. Kurzweil úgy véli, hogy a 2030-as évek elején eljutunk az úton. Ezen túl egyre drámaibb biotechnológiákat képzel: véresejt méretű "nanobotok" járőröznek testünkben, sejtszinten javítva a károsodást, sőt, visszafordítva az öregedési folyamatokat is. Azt javasolja, hogy végül kiválthatjuk a biológiai szerveket jobb szintetikus szervekre – például mesterséges véresejteket és nanotechnológiai tüdőket, amelyek hatékonyabban oxigéneltetik a vért, vagy nano-mérműi szívekkel, amelyek soha nem hibásodnak. Kurzweil szavaival élve: "végső soron a nanobotok képesek lesznek teljesen helyettesíteni a biológiai szerveket, ha szükség van rá vagy kívánsággal". Ennek eredményeként sok jelenleg halálos állapot gyógyíthatóvá vagy megelőhetővé válik. Kurzweil még azzal a provokatív állítással is egyetért, hogy "az első ember, aki ezerévig él, már megszületett" – nem egyetlen varázslatos miatt, hanem a génterápiák, sejtmegeújítás és nanotechnológiai javítások kombinációja miatt, amelyek folyamatosan meghosszabbítják az életet. Ez a fejezet tehát a közelebbi biotechnológiai csodákat mutatja be, az MI által tervezett gyógyszerektől és a CRISPR génszerkesztéstől kezdve a szervek regenerációján és az agy-gép integráción át az egészség kezelésében. Kurzweil szerint 2045-re maga az öregedés nagyrészt legyőzhető, és azok az emberek, akik szeretnék, végtelenül élhetnek fiatal, felnőtt testben.

7. fejezet: Veszély

Miután dicsérte az exponenciális technológia ígérteit, Kurzweil a 7. fejezetet az ilyen technológiák veszélyeinek és egzisztenciális kockázatainak szenteli. Fenyegetések spektrumát azonosítja, és azokat "valódi" veszélyekre és "észlelt" veszélyekre bontja (utóbbi nagyrészt a változástól való félelemből fakad). A valóságban Kurzweil őszinte: erősebb technológiát lehet szörnyű célokra használni. Megjegyzi, hogy az MI-vel működő autonóm fegyverrendszereket már állítólag bevetették a háborúban (például az orosz–ukrajna konfliktusban). Olyan világvégítéletes forgatókönyveket részletez, mint az MI által tervezett biológiai fegyverek: kemény, mindenevő "baktériumok", amelyek kicsúszhatnak az irányítás alól és "túlszárnyalhatnak a valódi baktériumokkal", vagy egy rosszul programozott MI, amelynek feladata egy vírus megölése, amely egy apró szemantikai hiba miatt "megöli az egészséges géneket". Komoly aggodalomnak tartja a nukleáris terjedést és a hiperszonikus rakétákat, valamint a mesterséges kórokozók véletlen laboratóriumi kibocsátását a következő évtizedekben. Ami a lényeges, Kurzweil nem állítja, hogy egyetlen megoldás lenne ezekre a sokféle kockázatokra. Ehelyett azt állítja, hogy szükségünk van egy mindent átfogó stratégiára – egy "koktélnak" a technológiai garanciákból, etikai elvekből és nemzetközi kormányzásból. Egy feltűnő pont, hogy maga az MI is része kell legyen a megoldásnak: például az MI használata az AI megfigyelésére és ellenőrzésére, vagy "versengő MI-k" alkalmazása a hibák észlelésére és egymás fékére. Olyan megközelítéseket ír le, mint az iterált erősítés, ahol egyre okosabb MI-ket felügyelt módon használnak, hogy biztosítsák a jövőbeli még erősebb MI-k összehangolását. Kurzweil aktív részt vesz az MI biztonságról szóló vitákban (segített kidolgozni a 2017-es Asilomar AI Elveket a biztonságos AI fejlesztéshez), és ebben a fejezetben ismételteti az "etikai bástyákra" és erős normákra a mesterséges intelligencia és biotechnológia visszaélésének megakadályozására. Másrészt Kurzweil figyelmeztet a

túlreagálásra, a haladás megállítására való visszaesésre. Kritizálja azt, amit "*fundamentalista humanizmusnak*" nevez – egy kialakuló neo-luddita attitűdöt, amely félelemből széles körű lemondást vagy technológiák betiltását követeli. Például említi, hogy a genetikailag módosított élelmiszerek elleni téves tiltakozások fokozták az éhezést Afrikában azzal, hogy akadályozták a segélyeket. Hasonlóképpen, megjegyzi, hogy néhányan még elvileg ellenzik az életmentő fejlesztéseket, mint a génterápia vagy a *fehérjehajtogató mesterséges intelligencia rákkutatásban*. Kurzweil álláspontja az, hogy nem szabad feladni a technológiai fejlődést, mert éppen ezek a fejlesztések képesek "*kezeln* az emberi szenvedést" – ehelyett felelősségteljesen kell haladnunk előre. A fejezetet óvatosan optimista hangon zárja: az emberiség már korábban is szembesült egzisztenciális veszélyekkel (emlékszik arra, hogy a látszólag elkerülhetetlen nukleáris háború árnyékában nőtt fel), mégis "*fajunk megtalálta a bölcsességet, hogy ne használjuk ezeket a szörnyű fegyvereket*". Ugyanígy úgy véli, hogy az MI-t, a nanotechnológiát és a biotechnológiát jóra tudjuk és kellene hasznosítani, miközben elkerüljük a katasztrófát. Az egzisztenciális fenyegetések valóságok, de bölcsességgel és éberséggel Kurzweil azt állítja: "*Nem vagyunk elítélve a kudarcra ezeknek a veszélyeknek a kezelésében.*".

8. fejezet: Párbeszéd Cassandrával

Az utolsó fejezetben Kurzweil egy kreatív szókratikus stílusú párbeszédet mutat be maga és egy képzelt szkeptikus, "Cassandra" között. Cassandra kifejezi azokat a kételyeket és félelmeket, amelyeket sok kritikus felvetett Kurzweil víziójával kapcsolatban, lehetővé téve számára, hogy pontonként reagáljon. Ezen formátumon keresztül Kurzweil közvetlenül a korábbi munkásságának és a szingularitás fogalmának kritikáit tárgyalja. Cassandra találoan aggódik a katasztrófa előrelátása miatt. Kihívást jelent Kurzweilnek olyan kérdésekben, mint például: Az emberi irrelevánság – "*Ha okosabbakat hozunk létre nálunk, számítanak-e még az emberek?*"; Jelentésvesztés – "*Ha a gépek végeznek minden munkát, sőt gondolkodnak helyettünk, mit csinálnak az emberek egész nap? Mi ad értelmet az életnek?*"; Egyenlőtlenség – "*Nem csak a gazdagoknak szólnak ezek a fejlesztések, disztópiát teremtve?*"; Istenként játszom – "*Elég bölcssek vagyunk ahhoz, hogy újrataláljuk az életet és az intelligenciát, vagy csak saját pusztításunkat okozzuk?*"; és Unmet előrejelzések – "*Mi a helyzet azokkal a dolgokkal, amelyek nem váltak ki a korábbi előrejelzéseidből?*". A párbeszédben Kurzweil szisztematikusan cáfolja ezeket az aggodalmakat védjegyének számító optimizmusával és adataival. Például, amikor Cassandra aggódik, hogy a szuperintelligens MI uralhatja vagy helyettesítheti az embereket, Kurzweil azt állítja, hogy "*a számítógépek nem versenyeznek velünk. Ők a kiterjesztésünk, elkísérnek minket az utunkon.*" Ragaszkodik hozzá, hogy az MI-vel való egyesülés inkább felhatalmazza az embereket, mint az emberiséget – 2045-re minden ember kreatív értelme az MI-hez való kapcsolódásokkal fog felerősödni, "*megnyitva egy végtelen bölcsesség és potenciál világát*", nem pedig egy Terminátor-stílusú átvétel. A jelentéskérdésre Kurzweil azt sugallja, hogy az emberi cél fejlődni fog; Ahogy ma sokan találnak beteljesülést a kreatív, tudományos és szabadidős tevékenységekben, ha az alapvető túlélési szükségleteket kielégítik, a jövőben "*milliószor több intelligenciánk*" lesz, amit művészetre, felfedezésre, személyes fejlődésre és mélyebb kozmikus rejtvények megoldására alkalmazhatunk. Az egyenlőtlenségről Kurzweil rámutat, hogy a fejlett technológiák kezdetben drágák és korlátozottak – mint például a mobiltelefonok vagy számítógépek –, de idővel olcsóvá és mindenütt elterjedté válnak, mindenki javát szolgálva. Úgy véli, ugyanez fog történni az MI fejlesztésekkel és a hosszú élettartamú kezelésekkel is: a korai befogadók lehetnek gazdagok, de "*ez a probléma idővel elmúlik*", ahogy a költségek zuhannak. Az egzisztenciális kockázatokkal kapcsolatban Kurzweil elismeri Cassandra félelmét (amelyek közül sok tükrözi a "Veszély" fejezetet), de ellenezi, hogy a *technológia teljes* lemondása sem megvalósítható, sem bölcs dolog. Ehelyett az aktív menedzsmentet hangsúlyozza: mesterséges intelligenciát

kell építenünk emberhez hasonló értékekkel, felügyeletet kell bevezetnünk, és fejlett MI-t kell használni a fenyegetések észlelésére és ellensúlyozására (például olyan MI rendszereket, amelyek éberen megakadályozzák a szökött MI-t abban, hogy a gemkapozó-maximalizáló forgatókönyvet követje). A párbeszéd során Kurzweil hangja türelmes magyarázatot tükröz, és magabiztos abban, hogy a világ "Cassandras" tévednek a komor ügyében. Valóban, ez a fejezet egy kritikusnak nevezett *"tanulmányozott válaszként szolgál a Cassandrák káosz előrejelzésére"* az MI és a jövő körül. A végén Kurzweil megismételi alaptézisét: hogy az emberi kreativitás és a technológia együttesen bármilyen kihívást képes legyőzni. Messze attól, hogy elveszítenénk emberi mivoltunkat, *"lényegében újraalkotjuk magunkat"* a jobb irányba ezekkel a fejlődésekkel. A Szingularitás szerinte nem végzet, hanem a következő evolúciós lépés – olyat, amelyhez nyitott szemmel kell közelíteni a kockázatokra, de izgalommal, nem kétségbeeséssel.

Kulcsfontosságú előrejelzések és technológiai előrejelzések

Kurzweil merész előrejelzéseiről ismert, és a *The Singularity Is Nearer* számos konkrét előrejelzést kínál az MI-ről és más technológiákról. A könyv néhány kulcsfontosságú előrejelzése és idővonala a következők:

- 2029 – Emberi szintű MI elérve: Kurzweil híresen megjósolja, hogy 2029-re az MI *átmegy a Turing-teszten*, és emberi szintű intelligencia lesz a legtöbb területen. Azt várja, hogy addig az MI képes lesz megkülönböztethetetlenül beszélgetni egy embertől, és a legjobb emberi szakértőkkel is teljesíteni szinte minden területen (ez az ő definíciója a mesterséges általános intelligencia, vagyis az AGI). Figyelemre méltó, hogy Kurzweil évtizedek óta teszi ezt a 2029-es jóslatot, és ragaszkodik ehhez ebben a könyvben – még azt is megfigyelve, hogy olyan fejlemények, mint a nagy nyelvi modellek, közelebb hoztak minket, odáig, hogy 2029 most már néhány megfigyelő számára szinte konzervatívnak tűnik.
- 2030-as évek eleje – Élettartam áttörései: Kurzweil azt jósolja, hogy a 2030-as évek elején elérjük a *"hosszú távú menekülési sebességet"*, ami azt jelenti, hogy az orvosi fejlődés gyorsabban meghosszabbítja az élettartamot, mint az idő múlik. A gyakorlatban ez azt jelenti, hogy minden évben a tudomány több mint egy év várható élettartamot ad, így az öregedés okozta halál már elkerülhetetlen. Kurzweil arra számít, hogy a 2030-as évekre orvosi nanobotok emberi kísérletekben vagy használatban lesznek – apró robotok, amelyek képesek bejutni a véráramba, hogy molekuláris szinten javítsák a sejteket és szerveket. Ezek lehetővé tennék, hogy az emberek végtelenül egészségesek maradjanak, kivéve baleseteket. Azt is előrelátja, hogy laboratóriumban termesztett szervek és szövetek széles körben elérhetővé válnak a 2030-as években, valamint génterápiákat, amelyek számos betegséget megszüntetnek. Kurzweil szerint a 2030-as évek végére maga a *halál nagyrészt opcionálissá válhat*, mivel meglesznek az eszközeink a testünk folyamatos javításához és megújításához.
- 2030-as évek – Agy-felhő interfészek: Egy fő előrejelzés, hogy a 2030-as évekre közvetlenül összekapcsoljuk az agyunkat a felhőalapú MI-vel. Kurzweil azt képzei, hogy nanobotok utaznak a kapillárisunkon keresztül az agyba, hogy integrálják neokérgenket a felhő gyakorlatilag korlátlan számítási kapacitásával. Ez gyakorlatilag összeolvotta az emberi és gépi intelligenciát. Ezt úgy írja le, mint hogy az elménk *képességei "milliódszor"* bővítik – azonnal hozzáférhetünk a tudáshoz és a kognitív képességekhez, amelyek jóval túlmutatnak a természetes agyunk határain. A 2030-as évek végére azok az emberek, akik ezt a technológiát használják, sokkal gyorsabban és összetettebben gondolkodhattak, mint a mai emberek. A mindennapi tapasztalatok

radikálisan megváltozhatnak: például bármely tény felidézése vagy egy összetett probléma megoldása azonnali lehetne az agy MI-vel való kapcsolatán keresztül. Kurzweil ezt az MI-vel való együttműködésnek nevezi, és úgy látja, hogy *ez egy együttműködés* módja az MI-vel, ahelyett, hogy versengünk vele.

- 2030-as évek – Széles körű gazdasági felfordulás és UBI: A technikai mérföldkövek mellett Kurzweil társadalmi és gazdasági változásokat is előrevetít. A 2020-as és 2030-as években jelentős munkaerő-piaci zavarokra számít az AI automatizálás miatt (ahogy az 5. fejezetben részletezzük). *A 2030-as évek közepére* a vezetés, gyártás és más területeken milliókat az MI és a robotok fognak kezelni. Erre válaszul Kurzweil azt jósolja, hogy a kormányok a 2030-as évektől kezdik el az univerzális alapjövedelem-programokat a menekültek támogatására. A korai UBI talán szerény, de mivel az MI által vezérelt termelékenység nagy vagyont teremt és csökkenti a költségeket, a társadalmi biztonsági hálók bővülni fognak. Kurzweil úgy véli, hogy a 2030-as évek végére az UBI vagy hasonló intézkedések elég erőssé válnak ahhoz, hogy mindenki profitáljon az AI jólétéből, ne csak a technológia tulajdonosai. Emellett reformokat tervez az oktatásban és a munkaerő-képzésben, hogy segítsék az embereket új szerepkörbe való átmenetbe az MI mellett dolgozva.
- 2040-es évek – Elmefeltöltés és digitális "halhatatlanság": Kurzweil szerint a 2040-es években a technológia lehetővé teszi a "túlvilági" lehetőségeket az elme feltöltésével. Azt jósolja, hogy körülbelül 2045-re (ami egybeesik a Szingularitással) az emberek képesek *lesznek "feltölteni"* elméjüket – vagyis létrehozni az agyuk és tudatuk teljes digitális emulációját. Ez lehetővé teszi a digitális halhatatlanság egyformáját: például, ha valaki biológiai teste meghal, egy tárolt elmefájl újrainstanciózható egy új alapon, például egy humanoid robotban vagy virtuális környezetben. Kurzweil még a "replikánsok" lehetőségét is említette – MI-alapú példányokat azoknak az egyéneknek, amelyek megőrzik személyiségüket és emlékeiket. A 2040-es évekre – írja – rendszeresen visszatartathatjuk az agyunk tartalmát, ahogy ma is az adatokkal. Ez mély kérdéseket vet fel (amelyeket a 3. fejezetben és a Cassandra párbeszédben vizsgálunk), hogy a másolat "te" vagy-e, és hogyan kezeli a társadalom több személyverziót vagy egy halott visszatérését. Ennek ellenére Kurzweil ezt a korlátok leküzdésére irányuló törekvésünk kiterjesztéseként tekinti: ahogy az orvosi fejlődés legyőzi a biológiai halált, az elmefeltöltés legyőzi az információs halált, lehetővé téve, hogy a tudatosság helyreálljon vagy szabadon utazzon testről testre.
- 2045 – A Szingularitás: Kurzweil továbbra is 2045-öt jelöli a Szingularitás évének – ezt a dátumot először a 2005-ös könyvében jósolta meg, és most is megerősíti. 2045 várhatóan az emberi és gépi intelligencia teljes összeolvadását jelenti, és egy olyan pontot, amelytől túl a változás sebessége szinte végtelenné válik (innen ered a "szingularitás" kifejezés). 2045-re a gépi intelligencia milliárdszor erősebb lesz, mint a segítség nélküli emberi intelligencia, lényegében "milliószoros" bővítés a közös agykapacitásunkban. Kurzweil azt jósolja, hogy addigra az MI nemcsak felgyorsul, hanem messze megelőzi az emberi agyat, és mivel összeolvadunk vele, *mi* is lényegében sokkal intelligensebbek leszünk. Gyakorlati szempontból Kurzweil egy majdnem utópikus 2045 utáni víziót fest: minden anyagi szükséglet kielégíthető nanotechnológiával és MI-vel (ami energia-, élelmiszer- és árubőséghez vezet); betegségek és öregedés gyógyulnak; a tudás és a kreativitás virágzik, ahogy mindenki hozzáférhet a szuperintelligenciához; És még a tudatos tapasztalatunk is mélyül, ahogy új művészeti formákat, tudományt fedezünk fel, és talán az idegen intelligenciával is kapcsolatba lépünk, miközben elterülünk a kozmoszban. Ez egy radikális átalakulás, amit Kurzweil elismer, *hogy "nehéz elképzelni"* teljesen. Gyakran használ metaforákat: az egyik, *hogy "olyan lesz, mintha egy barlanglakót kérnék meg, hogy értsen meg egy*

modern várost" – a Szingularitás utáni világ számunkra olyan, mint a mi világunk az őskori embereknek. Ennek ellenére ragaszkodik hozzá, hogy a lényeg *nem* idegenek inváziója vagy emberi kihalása, hanem az, hogy az emberek *valami* újabbá és vitathatatlanul csodálatossá fejlődnek. Ahogy tömören fogalmaz: "2045-re megteszünk a következő lépést az evolúciónkban. Képzeld el minden ember kreativitását a bolygón, ami összefügg a leggyorsabb számítógép sebességével és ügyességével... Ez a Szingularitás."

Ezek a jóslatok természetesen attól függenek, hogy Kurzweil exponenciális növekedési modelljei beálljanak. Olyan trendekre alapozza őket, mint a Moore-törvény (a számítási teljesítményre) és hasonló MI képességei, biotechnológiai költség-teljesítménye és hasonló pályáira. A könyv tele van exponenciális görbéket bemutató grafikonokkal, valamint történelmi adatokkal a történelmi adatok alátámasztják ezeket az idővonalakat. Kurzweil gyakran megjegyzi, amikor más szakértők egyetértenek egyetértésre az egykor vitatott előrejelzéseivel – például felidéri, hogy 1999-ben a legtöbb MI tudós azt hitte, hogy az emberi szintű MI *több száz* évre van, míg most, a 2020-as években sokan úgy vélik, hogy ez már a sarkon van (összhangban az eredeti 30 éves becsléséhez).

Kurzweil gondolkodásának fejlődése 2005 óta

Közel két évtized telt el Kurzweil korábbi könyve, *a The Singularity Is Near* 2005-ös megjelenése óta. *A szingularitás közelebb* című művében Kurzweil visszatekint arra, hogyan változott meg a világ és hogyan változott (vagy nem változott) saját gondolkodása:

- Alapvető előrejelzések változatlanok, most már mainstreamebbek: Kurzweil alapvető idővonala – az MI eléri az emberi szintet 2029-re, míg a Szingularitás 2045-re – *változatlan* marad. Ha valami, most már magabiztosabb ezekben a randikon. Megjegyzi, hogy ami 2005-ben rendkívül optimistának hangzott, egyre inkább elfogadják a technológiai közösségben. Például egykor kigúnyolt 2029-es MI-jóslatát néhány vezető visszhangozta vagy akár felülírta (ironian rámutat, hogy Elon Musk most azt mondja, hogy az AGI a 2020-as évek közepére érkezhetsen). Kurzweil a mélytanulás megjelenését, különösen a nagy nyelvi modellek (például GPT) közelmúltbeli áttöréseit igazolja annak a nézetének, hogy évtizedeken belül, nem évszázadokon belül drámai MI képességeket hozna létre. Interjúkban Kurzweilt megkérdezték, hogy akarja-e azt mondani: "Én mondtam" – megjegyzi, hogy valóban sokan, akik egykor azt hitték, hogy az AGI a távoli jövőben van, most már közelgőnek látják. Ez az elismerés, ha valami, megerősítette hitét a Gyorsító Hozamok Törvényébe.
- Új bizonyítékok és példák: 2005 óta számtalan technológiai előrelépés történt, és Kurzweil ezeket a könyv során beépíti. A gondolkodásmódja nem annyira változott, mint inkább *friss adatokkal frissítette*. Például 2005-ben Kurzweil általánosságban az MI fejlődéséről beszélt; 2024-ben kiemelheti az AlphaGo győzelmeit, önvezető autó prototípusokat, a GPT-4 képességeit és más konkrét mérföldköveket, amelyek bekövetkeztek. A biotechnológiában ami 2005-ben spekulatív volt (mint például a CRISPR génszerkesztés vagy az mRNA vakcinák), most valósággá vált, ezért Kurzweil ezeket használja arra, hogy megerősítse érvét, miszerint az általa jóslott pályán haladunk. Röviden: *a The Singularity Is Nearer* kevésbé új elméletek bemutatásáról szól, inkább arról, hogy megmutassa, hogyan illeszkedik a mai élvonalbeli technológia Kurzweil meglévő keretrendszeréhez. Egy kritikus megjegyezte, hogy a folytatás "a Szingularitás közel alapötleteit veszi fel, és a legújabb technológiai fejlesztésekkel erősíti őket." Kurzweil számos fogalma – exponenciális növekedés, hat korszak, elme feltöltés – már

2005-ben is megvitatták; most pedig a jelenlegi kontextusban és azzal az érzéssel, hogy közelebb vagyunk a beteljesülésükhöz (innen a "Nearer").

- A szkeptikusok és nehéz kérdések kezelése: A könyv egyik figyelemre méltó fejlődése Kurzweil közvetlenebb kritikákhoz való bekapcsolódása. 2005-ben, bár a kockázatokról beszélt, Kurzweilt gyakran kritizálták, amiért *túl optimista* és elsiklik a társadalmi összetettséget. A 2024-es könyvben egész fejezeteket szentel a munkahelyeknek és az egzisztenciális kockázatoknak ("Peril"), a záró párbeszéd pedig kifejezetten szkeptikus nézőpontokat érint. Ez arra utal, hogy Kurzweil felismerte, hogy gyakorlati és etikai alapon kell meggyőzni az olvasókat, nem csak egy rózsás jövőt festeni. Például most hangsúlyozza a technológia intézményi elterjedésének lemaradását: elismeri, hogy miközben maga a technológia exponenciálisan fejlődik, az emberi egyének, különösen az intézmények lassabban alkalmazkodnak. A *The Singularity Is Nearer* című művében elismeri, hogy ez a páratlanság késleltetheti az innovációk gyakorlati bevezetését – utalás arra, hogy néhány korábbi idővonala bizonyos alkalmazásokra (például a teljesen önvezető autók 2010-es évekre) túl agresszív lehetett. Azonban, ahogy szokásához hasonlóan, gyakran azt állítja, hogy ezek az akadályok végül leomlanak, mert az előnyök túl nagyok ahhoz, hogy ellenálljanak (például *"az orvosi nanobotok elleni ellenállás eltűnik, mert a betegség és a halál olyan szörnyű"*). Így bár Kurzweil megoldásai optimisták maradnak, ő egyértelműbb a súrlódásokról és akadályokról az út során.
- Finomított fogalmak és terminológia: Kurzweil definíciói egyes területeken élesebbek voltak. Például tisztázza, mit ért "emberi szintű MI" és "AGI" alatt. Interjúkban elmagyarázta, hogy 2029-re olyan MI-k lesznek, amelyek felvethetők az *emberek legjobb képességeivel bizonyos területeken*, és röviddel ezután olyan MI, amely mindent meg tud csinálni, amit egy ember képes (amit ő AGI-nek is nevez). Magas mércét állít az AGI-nak – *"a legjobb emberek képességei minden területen"*, nem csak átlagos emberi képesség –, ezért nem mozgatta el a dátumát előre, annak ellenére, hogy gyors előrehaladás. Emellett időt szán arra, hogy megkülönböztesse a *szűk* MI-t az *általános* MI-től, és hogy a jelenlegi GPT modellek, bár lenyűgözőek, mégis vannak ismert hiányosságai (kontextushossz, józan ész stb.), amelyeket nagyobb számítási kapacitással és adattal lehet leküzdeni. Emellett Kurzweil mindig is nem kedvelte a "mesterséges" intelligencia kifejezést (mivel a gépi intelligenciát valódi intelligenciának tekinti). Ebben a könyvben ismételteti ezt az álláspontot, és megpróbálja az MI-t a természetes intelligencia kontinuumjaként beállítani, nem valami "másként" – így az MI-vel való egyesülés a saját intelligenciánk folytatása, nem idegen invázió.
- Nagyobb hangsúly az etikai keretekre: Bár Kurzweil továbbra is alapvetően optimista, 2005-nél nagyobb hangsúlyt fektetett az etikára és a kormányzásra. Megemlíti részvételét olyan kezdeményezésekben, mint az Asilomar AI Elvek, és nemzetközi együttműködést szorgalmaz az MI biztonság terén. 2005-ben a főszereplő narratíva az ígéret a veszély helyett; 2024-ben Kurzweil továbbra is az ígérethez hajlik, de több oldalt szentel arra, hogy biztosítsa az olvasókat, komolyan veszi a fenyegetéseket, és megvitatja, hogyan kezelje őket. Ezt úgy lehet feldolgozni, hogy a közvélemény növekvő aggodalma az MI iránt (amely ma sokkal nagyobb, mint 2005-ben). Ezért most a gondolkodásmódja egy sürgetőbb felelősségteljes vezetési igényt foglal magában, hogy a technológia hasznos legyen. Kurzweil alapvető nézete azonban – miszerint *a haladás betiltása vagy lassítása nem a megoldás* – nem változott. Ha valami, most már történelmi analógiákkal (például hogyan kerülte el a társadalom a nukleáris háborút, vagy hogyan buktak meg a luddita lázadások), hogy megerősítsék azt az érvet, hogy bölcsességgel kell szembenéznünk a jövővel, nem pedig félelemmel.

Összefoglalva, Kurzweil átfogó víziója a *Szingularitás közelebb című* műben nagyon is összhangban áll a 2005-ös víziójával. Ami fejlődött, az a kontextus: az MI most már a címlapos hír, sok jóslatai részben valóra vált, és finomítja a narratívát, hogy megválaszolja a "Igen, de..." kérdések. Az optimizmus továbbra sem halványulhat – ha valami, Kurzweil most már jobban érzi magát – de egyre nagyobb hajlandóságot mutat arra, hogy foglalkozzon az ellenvéleményekkel, és részletezze az átmeneti lépéseket és védelmi intézkedéseket az úton 2045 felé.

A könyv kulcsfontosságú technológiáinak ábrázolása

Kurzweil a *The Singularity Is Nearer* című művében *széles körű technológiákat dolgoz fel*, gyakran a "GNR" (Genetics, Nanotechnology, Robotics) rövidítés alá sorolva, az MI pedig a robotika és szoftverek egyik mozgatórugója. Íme, hogyan mutatják be és vetítik be a jelenlegi technológiákat:

- **Mesterséges intelligencia (MI):** Az MI a könyv szakdolgozatának középpontjában áll. Kurzweil az MI-t úgy írja le, mint *"gyorsan fejlődik az emberi képességeken túl"* egyik területen a másikig. Áttekinti, hogyan képes a modern MI már most is legyőzni az embereket összetett játékokban, bizonyos orvosi képeket pontosabban diagnosztizálni, mint az orvosok, zenét komponálni, sőt koherens szöveget is generálni. 2029-re állítja, hogy az MI szinte minden területen felülmúlja az emberi intelligenciát. Kurzweil egy fontos fogalommal rendelkezik a gyorsító hozamok törvénye az AI kontextusában: minden fejlesztés az MI-ben (pl. jobb hardver vagy algoritmusok) gyorsabban vezet a következő fejlődéshez, exponenciális növekedési görbét hozva létre. Például a nagyobb és gyorsabb számítógépek tették lehetővé a mélytanulási forradalmat, amely ma még jobb chipek és szoftverek tervezésében segít, és így tovább. Kurzweil a jelenlegi MI kihívásokkal is foglalkozik: megjegyzi, hogy a mai modellek néha *"hallucinálnak"* vagy hiányzik a józan ész, de azt várja, hogy ezeket a problémákat jóval 2029 előtt megoldják a megnövelt számítási kapacitással és jobb képzéssel (rámutatva, hogy a GPT-4 már sokkal kevesebbet hallucinált, mint a GPT-2, stb.). Hangsúlyozza, hogy az MI valódi intelligencia – elutasítja azt az elképzelést, hogy "hamis" lenne, csak azért, mert szilícium alapú –, és úgy véli, hogy a fejlett MI valamilyen tudatosságot elér (amit a 3. fejezetben érvel). Végül soron az MI-t úgy ábrázolják, mint *eszközt, amely mindent átalakít*: kreatív partnerünk lesz (segít új gyógymódokat, művészeti formákat, tudományos elméleteket feltalálni), a tanárunk (aki minden gyereket személyesen korrepetál), a dolgozónk (a rutinmunka automatizálása), és végül *még önmagunk egy része is* (amikor az agyunkat összekapcsoljuk vele). Kurzweil MI-ről alkotott nézete túlnyomórészt pozitív – ő úgy látja, hogy *"az emberiség kiegészítője"*, nem pedig helyettesíti. A veszélyeket elismerik, de ő kijelenti, hogy a megfelelően összehangolt MI *lesz "szövetségesünk, nem ellenségünk."*
- **Agy-számítógép interfészek (BCI):** A könyv egyik feltűnő előrejelzése az idegi interfészek fejlesztése, amelyek integrálják az emberi agyat az MI-vel. Kurzweil *nanobotokat* (nano-méretű robotokat) lát előre, amelyek nem invazívan képesek bejutni a véráramba és az agyba jutni. Ott ezek a nanobotok vezeték nélküli kapcsolatot hoznak létre a neuronjaink és a felhő között. Kurzweil azt írja, hogy a 2030-as évekre ez a technológia lehetővé teszi, hogy *a "neokéreg felső rétege"* közvetlenül összekapcsolódjon a felhőalapú MI rendszerekkel. Gyakorlatilag ez azt jelenti, hogy a biológiai agyad zökkenőmentesen hozzáférhet a szuperemberi számítási kapacitáshoz és memóriához. Lehetett "gondolkodni" egy kérdésen, és a válasz a felhőkapcsolaton keresztül alakult ki az elmédben, ahogy most szóban kérdezzük a Siri-t vagy a Google-

t (de sokkal gyorsabban és integráltabban). Kurzweil ezt a folyamatot úgy írja le, mint *"megosztani a neokortexünket"* az MI-vel – gyakorlatilag egy hibrid intelligencia közös létrehozásával. Idővel a biológiai neuronokból érkező gondolatok aránya a felhőalapú MI-hez képest egyre inkább az AI oldalára tolhat el, különösen, ahogy az emberek egyre több agyimplantátumot vagy neurális nanorobotot alkalmaznak. 2045-re Kurzweil arra számít, hogy a BCI-k lehetővé teszik az elme és gép teljes összeolvasását, ahol nehéz megmondani, hol ér véget "te" és hol kezdődik a felhő. A könyv megnyugtatja az olvasókat, hogy ez a helyzet *önkéntes és fokozatos*: az emberek az agyfelületeket fogják használni az egyértelmű előnyök (fejlett intelligencia, kommunikáció, memória) miatt, és mert ennek elmulasztása olyan lenne, mintha *"megtagodnánk az okostelefonok használatát"* a mai világban – technikailag egy lehetőség, de ez messze lemarad. Kurzweil a folyamatban lévő fejleményeket is bizonyítékként említi: már léteznek primitív BCI-k (például agyi implantátumok, amelyek lehetővé teszik a bénult betegek robotkarját gondolataikkal, vagy Elon Musk Neuralink kísérletei). Ezek kezdetlegeseek, de megmutatják a megvalósíthatóságot. Kurzweil BCI víziója lényegében a Szingularitás kapuja – így *egyesülünk* az MI-vel. És úgy véli, hogy kevésbé lesz idegen, mint amilyeneket tartanak: *"Gondolj rá úgy, mintha az okostelefonod az agyadban lenne"* – mondta egy interjúztató – nem önmagad elvesztésként fogod tapasztalni, csak *a képességeid extrém fejlődését*.

- Genetika és biotechnológia: A "Genetika" széles címkéje alatt Kurzweil magában foglalja a génszerkesztést, a biotechnológiát és az élettudományi innovációkat. Ezt forradalomként ábrázolja, ugyanolyan fontos, mint az MI. A könyv a CRISPR-ről és más génszerkesztő eszközökről beszél, amelyek lehetővé teszik a DNS pontos módosítását. Kurzweil azt jósolja, hogy a legtöbb genetikai betegséget a rossz gének kivágásával gyógyítjuk meg (például olyan gének kizárásával (például olyan gének kizárása, amelyek olyan betegségeket okoznak, mint a Huntington-szindróma, a BRCA-rák mutációk stb.). De a betegségek gyógyításán túl előrelátja, hogy genetikát alkalmaznak az emberek fejlesztésére – potenciálisan erősítve az immunrendszerünket, javítva az anyagcserét, sőt, még olyan tulajdonságokat is finomíthat, mint az intelligencia vagy a személyiség (bár enyhén érinti a tervező babák etikai aknamezőjét). Egy másik fókusz a biotechnológia és az MI konvergenciája: Kurzweil az MI-vezérelt kutatásról beszél a gyógyszerkutatásban és az orvosi kezelésben (ahogy az MI és gyógyszerkutatásról szóló részletben látható). A kezelések automatizálásával az MI képes "billiónyi molekulát" átszűrni, hogy órák alatt találjon gyógymódot, ami gyors új gyógyszerek és vakcinák fejlesztéséhez vezet. Kiemeli a szintetikus biológia fejlődését is – a sejtek programozása, mint a szoftver –, amelyek lehetővé tennék szövetek vagy akár szervek újranolesztését. Például megfigyelhetők a sejtek fiatalosításra történő újraprogramozását (például néhány öregedésellenes kutatásban). Kurzweil azt jósolja, hogy néhány évtizeden belül a biotechnológia hatékonyan *újraprogramozza az öregedés biológiáját*: kikapcsolja a káros géneket, aktiválja a védő géneket, és helyreállítja a sejtkárosodást. Ez kulcsfontosságú a hosszú életre vonatkozó víziójához. A genetika/biotechnológia a *The Singularity Is Nearer* című műben alapvetően arról szól, hogy *"elsajátítsák a biológia információs folyamatait"*, ahogy Kurzweil máshol fogalmazott. A DNS egy kód, a sejtek biológiai gépek – és kezdjük feltörni ezt a kódot. Kurzweil jövőjében a génterápia szedése ugyanolyan rutinszerű lehet, mint egy tablettá szedése, és egyedi tervezésű vírusokat küldhetnek a szervezetbe, hogy menet alatt átalakítsák a sejteket (jó célokra, például a rák megsemmisítésére). Az összesített ábrázolás rendkívül reményteli: a betegséget és az öregedést mérnöki problémákként látjuk, amelyeket megoldunk, nem pedig az élet megváltoztathatatlan tényeiként.

- Nanotechnológia: A nanotechnológia – az anyag molekuláris és atomi szintű manipulálása – Kurzweil előrejelzésének másik pillére. A könyv a nanotechnológiát olyan területként írja le, amely *"forradalmasítja az orvostudományt és a gyártást"*. Kurzweil orvostudományban példákat hoz, például nanorobotokat, amelyek miniatűr immunrendszerként működnek, kórokozókat zúzva és sejtenként javítanak szöveteket. Híresen képzelte el a "respirocitákat" (nano mesterséges vörösvérsejteket), amelyek egy nap lehetővé tennék, hogy egy órán át visszatartsd a lélegzeted, vagy lélegzés nélkül futhass, mert olyan hatékonyan juttatják az oxigént. A gyártásban a molekuláris összeszerelőkről beszél – lényegében nanotechnológiai gépekről, amelyek atomonként képesek bármilyen anyagot vagy terméket atomonként felépíteni. Ez vezet a *"bőség" korszakához: ha olcsón tudod kezelni az anyagot atomi szinten, szinte bármilyen anyagi árut minimális költséggel lehet létrehozni*. Kurzweil azt javasolja, hogy a nanotechnológia rendkívül olcsóvá tehetné a napelemeket (tökéletes apró szerkezetek összeszerelésével), megtisztíthatja a szennyező anyagokat (nanogépeket, amelyek szemetet vagy olajszivárgásokat bontnak), sőt, akár élelmiszert vagy ruházatot is építhetnének alapvető kémiai alapanyagokból. Egy élénk jóslat: a testünkben lévő nanobotok folyamatosan belülről építhetnek minket, így gyakorlatilag halhatatlanná és ellenállóvá válhatunk a betegségekkel. Úgy gondolja, hogy a 2040-es évekre ha valaki súlyosan megsérül, nanorobot rajok képesek sejteket rekonstruálni és gyorsan gyógyítani a sérüléseket. A nanotechnológia kulcsfontosságú Kurzweil úrvíziójában is – rendkívül hatékony nano-méretű gyártás lehetővé tehetné olcsó űrlakóhelyeket vagy terraformáló technológiákat (bár a könyv nem mélyen merül el az űrben, inkább a földi kérdésekre koncentrálna). Fontos az, hogy Kurzweil felismeri a nanotechnológia veszélyeit is, például a hipotetikus "szürke nyálk" helyzetet, amikor az önreplikáló nanogépek szabadon futnak. A Peril fejezetben említi a *"nano-alapú fegyvereket"* és a szökött nanotechnológia kockázatát. De ismét a megelőző intézkedésekre tegyél, és azt állítja, hogy az előnyök (betegségek gyógyítása, a hiány megszüntetése) messze felülmúlják a kockázatokat, ha megfelelően kezelik őket. Összefoglalva, a nanotechnológiát a könyvben a végső eszköztárként ábrázolják – a fizikai világ legkisebb léptékű irányítása, amely lehetővé teszi számunkra, hogy megoldjuk a korábban megoldhatatlan fizikai problémákat (mint például betegségek, szennyezés, erőforrás-korlátok).

Ezen kívül Kurzweil érinti a robotikát is (fizikai robotok értelemében, bár gyakran használja a "robot" kifejezést mesterséges intelligencia szoftverek bevonására). Elvárja, hogy az intelligens robotok elvégzik a legtöbb veszélyes vagy ismétlődő munkát. Röviden beszél a kvantumszámításról is, csak hogy azt mondja, nem tartja annak szigorúnak a szingularitás eléréséhez – úgy véli, hogy a klasszikus számítástechnikákkal és a 3D chiparchitektúrákkal elérhetjük azt, amire szükség van. (Kissé szkeptikus a kvantumszámítás gyakorlati értékével kapcsolatban, ami érdekes lábjegyzet a technológiai értékeléseiben.) Egy másik említett technológia a virtual/augmentált valóság: Kurzweil azt javasolja, hogy ahogy közelítjük a szingularitást, a virtuális világok megkülönböztethetlenné válnak a valóságtól, és az emberek VR környezetekben töltenek időt, amelyek ugyanolyan jelentősek, mint a fizikai valóság. Ez összefügg az elme feltöltésével – végül a *"tudatosságod"* VR-ben működhet, ha úgy döntesz, elmosódva a fizikai és digitális létezés határát.

Összességében Kurzweil a jelenlegi technológiákat összefonódóként és gyorsulóként mutatja be. Az MI segíti a biotechnológia fejlesztését; a biotechnológia a nanotechnológiát fogja kihasználni; a nanotechnológia jobb AI hardvert épít és így tovább, pozitív visszacsatolási ciklusban. Ez az összekonvergencia az, ami a Szingularitás felé hajt minket. Minden

technológia nem elszigetelten, hanem az emberi fejlődés nagy narratívájának részeként jelenik meg: MI + BCI + genetika + nanotechnológia = túllépve korlátainkat.

Filozófiai és társadalmi következmények

A technikai előrejelzéseken túl a *Szingularitás közelebb* számos filozófiai és etikai kérdést vizsgál a jövőnkkel kapcsolatban. Kurzweil víziója mély kérdéseket vet fel arról, mit jelent embernek lenni, a halhatatlanság kereséséről, a tudatosság természetéről, és arról, hogyan fog a társadalom változni. A könyv különböző fejezetekben foglalkozik ezekkel (különösen a 3. fejezetben, a "Ki vagyok én?" és a Cassandra párbeszédben). A főbb következmények a következők:

- Halhatatlanság és radikális élethosszabbítás: Kurzweil egyik legvitatottabb álláspontja az, hogy a halál legyőzhető, vagy legalábbis határozatlan időre elhalasztható. A gyakorlati *halhatatlanság* eléréséről beszél a biomedikai és digitális eszközök kombinációjával. Filozófiailag ez megkérdőjelezi azt az ősi nézetet, miszerint a halandóság értelmet ad az életnek. Kurzweil azonban azok mellett áll, akik úgy vélik, hogy hosszabb (akár korlátlan) élet nettó pozitívum lenne. Híresen viccelődött, hogy személyes terve az, hogy *"elég sokáig élni ahhoz, hogy örökké éljünk"*, vagyis túlélni, amíg meg nem érkezik a halhatatlanság technológiája. A könyvben a hosszú élet menekülési sebességét tárgyalja (elég hosszú élet ahhoz, hogy tovább növelje az élethosszabbítást), és azt sugallja, hogy ma élők sokáig gyakorlatilag soha nem halnak meg öregkorban. Ez etikai kérdéseket vet fel: Ki élhet határozatlan időre? Hogyan fogják az erőforrásokat elosztani egy olyan világban, ahol az emberek évszázadokat élhetnek? Kurzweil röviden figyelembe veszi az erőforrás-korlátokat, de úgy véli, hogy a fejlett technológiák (mint például a nanotechnológia és az MI által kezelt gazdaságok) bőséget biztosítanak egy sokkal nagyobb, hosszabb életű népesség támogatására. Van egy pszichológiai dimenzió is: ha az elméket vissza lehet állítani és helyreállítani, az élet és halál határa elmosódik. Az elhunyt mesterséges intelligencia segítségével való "visszahozás" gondolata (pl. valakit az adataiból újraalkotni) a *2040-es évek "túlvilági"* technológiájaként említik. Kurzweil elismeri, hogy ez *"érdekes társadalmi és jogi kérdéseket"* fog felvetni – például, vajon egy személy digitális másolata ugyanazokkal a jogokkal rendelkezik, mint az eredeti, és mi történik, ha több példány létezik? Nem ad végleges válaszokat, de azt sugallja, hogy a társadalomnak új normákat kell kialakítania. Kurzweil általános hozzáállása az, hogy az élet értékes, és több élet jobb: ha választhatnának, a legtöbben tovább egészséges és életerő maradnának, így az élet meghosszabbítása erkölcsi kötelezettség. Még úgy is beállítja ezt, mint *"a sorsunkat a saját kezünkbe helyezni"*, ahelyett, hogy a sors vagy a biológia kegyelmére bízánk.
- Tudatosság és identitás: Talán a legmélyebb filozófiai kérdés az, mi lesz "velünk", amikor egyesülünk a technológiánkkal. Kurzweil funkcionista szemléletet képvisel a tudatosságról – hogy ha egy rendszer megkülönböztethetetlenül viselkedik egy tudatos embertől, akkor tudatos. Ez azt jelenti, hogy az elméd számítógépre való feltöltése vagy a gondolataidat megosztva egy MI-vel nem "öli meg" a tudatot; inkább a tudatosság nem biológiai alapokban is lebeghet. Elképzeli, hogy egy elméd másolata a felhőben valóban *úgy érezné magát, mint te*. Ez felveti a klasszikus identitáskérdést: ha két példányod létezik (egy biológiai, egy digitális), akkor mindkettő "te" vagy? Kurzweil nem merül el túl mélyen a személyes identitás paradoxonban, de inkább arra hajlik, hogy a minták (információ és folyamat) kontinuitása számít, nem az eredeti atomok kontinuitása. A 3. fejezetben, amelynek címe "Ki vagyok?" azon gondolkodik, hogy "te" mennyi információs mintázata az agyadban. Ha ezek a minták megőrizhetők vagy

kiterjeszhető, akkor kitartasz, még ha a közegben is változik. Van egy csoportos vagy univerzális tudatosság vizsgálata is: Kurzweil utal arra, hogy ahogy összekötjük az agyunkat és az MI-vel, kialakulhat egy globális agy vagy kollektív tudatosság. Utal arra az elképzelésre, hogy az univerzum felébredése (6. korszak) azt jelentheti, hogy minden anyag intelligenciával van átítva, ami szinte panteista vagy panpszichista vízió. Ez természetesen spirituális/filozófiai terület határát képezi – Kurzweil, bár nem vallásos, gyakran kvázi-spirituális metaforákat mond arról, hogy "az univerzum tudatossá válik önmagáról." Társadalmi szinten, ha az elméket le lehet másolni, fejleszteni és összeolvadni, az ember definíciója is megváltozhat. Az emberek több személyiséget is felveszhetnek (például különböző feladatokhoz futtathatnak másolatokat), vagy akár több ember *is összeolvadhat* az elméjüket egy közös élményhez. Ezeket a helyzeteket nem vizsgálják mélyen, de az a következtetés, hogy *az egyéni identitás sokkal folyékonyabbá válhat*. Kurzweil valóban odafigyel az *MI elmék etikai státuszára*: ha egy MI tudatos és jogkért könyörög, akkor Kurzweil álláspontja az, hogy ezt komolyan kell vennünk. Bevezeti az etikai fogalmát a nem emberi tudatosság számára, azzal érvelve, hogy fel kell készülnünk arra, hogy tisztelettel és erkölcsi megfontolással kezeljük az intelligens gépeket és a feltöltött elméket. Ez radikális elmozdulás a jelenlegi emberközpontú erkölcsünkhöz képest.

- Etika, felelősség és összehangolás: A nagy hatalommal nagy felelősség jár. Kurzweil hangsúlyozza az "erkölcsi kötelezettséget", hogy ezeket a hatalmas technológiákat jóra kövessük, miközben kontrolláljuk azok hátrányait. Etikailag az egyik fő vita az MI összehangolásáról szól – arról, hogy az AI céljai összhangban legyenek az emberi értékekkel, hogy a szuperintelligens MI előnyös legyen nekünk, és ne ártson akaratlanul. Kurzweil úgy véli, hogy ez elérhető, és valójában már most is az MI közösség fókuszja (*"Minden nagy vállalat több energiát fektet arra, hogy rendszereik biztonságosak legyenek és összhangban legyenek az emberi értékekkel, mintsem új fejlesztések létrehozására"* – jegyzi meg). Ellenzi az MI kutatás leállításának gondolatát, az ilyen átfogó ellenállást *"nem ésszerűnek"* nevezi, és összehasonlítja a korábbi félelmekkel az új, kezelhetőnek bizonyultó technológiák miatt. Ehelyett olyan normákat és esetleg szabályozásokat sürget, amelyek ösztönzik a biztonságos MI fejlesztését – például nemzetközi megállapodásokat az autonóm fegyverekről vagy az MI átláthatósági irányelveket. Kurzweil alapvetően optimista az emberi bölcsességgel kapcsolatban: a nukleáris háború példáját idézi, hogy megmutassa: az emberiség megfelelő előrelátással visszaléphet a szakadék széléről. Egy másik etikális/társadalmi kérdés az egyenlőtlenség és a hozzáférés. Kurzweil elismeri, hogy *lesz átmeneti egyenlőtlenség* (a gazdagok először kapják a technológiát), de erősen hiszi, hogy a technológiai trendek idővel demokratizálják a hozzáférést. Egy kapcsolódó kérdés, hogyan kerüljük el a techno-elit uralkodó osztályt – Kurzweil válasza az, hogy a technológia demokratizálódása felhatalmazza az egyéneket (például a 2030-as évekre egy gyerek laptopkal és MI-vel képes olyasmit tenni, amit évtizedekkel korábban csak a kormányok vagy vállalatok tudtak megtenni). Lényegében abban a történelmi trendben bízunk, hogy a technológia a sokakat erősíti, nem csak a keveseket, különösen, amikor olcsóvá és mindenütt elterjedté válik. Mégis, a könyv rendkívüli élettartamra és intelligenciára vonatkozó előrejelzései aggályokat vetnek fel azzal kapcsolatban, hogyan fog a társadalom megbirkózni a társadalmi struktúrák potenciálisan hatalmas változásaival: nyugdíj, család, karrier, sőt, még az alapvető dolgok, mint a munka jelentése vagy az emberi erőfeszítés értéke. Kurzweil utal rá, hogy *"meg kell változtatnunk a cél és jelentés meghatározásánkat"*, mivel a munka már nem szükséges a túléléshez. Utal Viktor Frankl elképzelésére, miszerint a jelentés elengedhetetlen az emberek számára, utalva arra, hogy a szűkös világ után a jelentést kreatív

törekvésekben, kapcsolatokban, tudásban stb. keressük, nem pedig pusztán megélhetésben.

- Tudatos MI és jogok: Kurzweil egy konkrét filozófiai kérdés az, hogy hogyan kezeljük a tudatos MI-t vagy a fejlett állatokat. A 3. fejezetben megjegyzi, hogy a társadalom régóta megtagadta az állati tudatosságot, hogy igazolja az állatok érzéketlen erőforrásként való kezelését. A fejlődés bizonyítékaként a Cambridge-i Tudatossági Nyilatkozatot (2012) idézi, amelyben az idegtudósok azt állították, hogy az emlősök, madarak és más élőlények valószínűleg tudatos élményeket élnek át. Párhuzamot vont a jövőbeli MI-vel: hajlamosak lehetünk megtagadni egy MI belső életét, még akkor is, ha az pontosan tudatos lényként viselkedik. Kurzweil egyértelműen nyitott gondolkodású álláspontot szorgalmaz – ha úgy jár és beszél, mint egy kacsa (tudatos lény összetettsége van), akkor feltételeznünk kell, hogy van elméje. Ez a jövőben jogi jogokat vagy személyiséget adhat MI entitásoknak, ami filozófusok és futuristák által vitatott téma. Kurzweil nem vázol fel jogi javaslatokat, de azzal, hogy felhossa ezt, azt jelzi, hogy a *"személy" definícióját bővíteni* kell. Implicit módon érinti az elmemódosítás etikáját is: ha képesek vagyunk növelni az intelligenciát, vajon kötelességünk van rá? Lehet-e a fejlesztések elutasítása egyenértékű például azzal, mintha ma nem adunk oktatást egy gyermeknek? A társadalmi normák úgy változhatnak, hogy a *"természetes" nem feljebb lábazott emberek ritkává válhatnak vagy akár hátrányos helyzetűnek* tekinthetők. Ezek spekulatívak, de Kurzweil jövője olyan lesz, ahol az emberiség rendkívül sokszínű – néhányan biológiai maradnak, mások kiborgok vagy teljesen digitális lények lehetnek. Ennek a sokszínűségnek a tolerálása és integrálása filozófiai és kulturális kihívás lesz.
- Emberi jelentés és spiritualitás: Bár Kurzweil technológus, munkái gyakran érintenek kvázi-spirituális témákat. Magát a Szingularitást olyan kifejezésekkel írják le, amelyeket egyesek vallási jóslathoz hasonlítanak – ezt még egy kommentátor is humorosan megjegyezte, aki azt mondta, Kurzweil Szingularitása *"közel vallási megvilágosodási esemény" elemeit* tartalmazza követői számára. Kurzweil maga metaforikus nyelvezetet használ a transzcendencia nyelvezetében: *"elménk fejlődése", hogy "mélyebb betekintést nyissunk fel" és "milliószor bővítsük intelligenciánkat"*. Van egy utópikus aljavonat, amely szerint ez majdnem mennyei állapotot hozhat (kifejezetten vallási keretek nélkül). Az általa képviselt elme filozófiája lényegében materialista (az elme mint számítás), de az általa elképzelt végső állapot – az univerzum felébredése – nagyszerű, szinte kozmikus sors ízű. Néha metaforikus módon utal Istenre vagy spiritualításra (például azt mondja, hogy a Szingularitás lehetővé tenné számunkra, hogy *"Istent játsszuk"*, saját világokat és életet teremtve, vagy elérjük azt, amit egyesek *"isteni"* megértést neveznek minden tudásról). A könyv finoman érinti a tudatosságot az univerzum céljának – ez az ötlet összhangban van Teilhard de Chardin Omega Point című művének egyes értelmezéseivel vagy más filozófiákkal, ahol az evolúciónak iránya van. Kurzweil gyakorlati álláspontja azonban az, hogy *mi teremtünk értelmet*. Az egyik idézett sora így szól: *"A jövőbe nem lépünk bele. A jövőt mi teremtjük."* – kiemelve az emberi cselekvési képességet. Ez egy alapvető humanista filozófiát tükröz: ahelyett, hogy a számunkra adott jelentést keressünk, bővülő képességeinket használjuk a jelentés formálására.

Összefoglalva, Kurzweil könyve mély kérdésekre ösztönzi az olvasókat: Ha örökké élhetnél, te tennéd? Mi történik, amikor az elméd bárhol és mindenhol lehet? Egy digitálisan feltöltött személy ugyanaz, mint az eredeti? Egyedül az intelligencia képes megoldani az emberiség erkölcsi dilemmáját, vagy inkább kockáztatjuk, hogy súlyosbítjuk hibáinkat? Kurzweil optimista, hogy ahogy okosabbak és összekapcsoltabbak leszünk, *bölcsőbbek és etikusabbak is*

leszünk, a technológiát értékeink emelésére használva, nem pedig megrontítani azokat. Nem mindenki ért egyet – a kritikusok attól tartanak, hogy az exponenciális hatalom ugyanolyan könnyen felerősítheti a kapzsiságot vagy a zsarnokságot. Kurzweil elismeri ezeket a félelmeket (például Cassandra hangján keresztül), de történelmi bizonyítékokkal válaszol rájuk pozitív társadalmi változásról, valamint az emberi alkalmazkodóképességbe és jószágba vetett hittel. A szeretet, a kreativitás és kíváncsiság – javasolja, még mindig hajtanak minket a Szingularitásban, ahogy korábban is – csak akkor lesz sokkal nagyobb lehetőségünk kifejezni őket.

Korábbi munkásságának kritikái

Ray Kurzweil régóta lelkes támogatót és határozott kritikát kap. *A szingularitás közelebb* című művében aktívan foglalkozik a korábbi jóslataival és filozófiáival szemben felmerült gyakori kritikákkal:

- "Nincs időzítésed / túl optimista vagy." Egy gyakori kritika, hogy Kurzweil idővonalai túlzottan optimisták vagy akár egyszerűen tévesek. A kritikusok gyakran említenek olyan példákat, mint például, hogy 2023-ra még nem voltak teljesen agyi kontrollált VR vagy nano-rajok, amelyek rutinként megjavították a rákot (néhány dolog utalt korábbi írásaiban a 2010–2020-as évekre). Ebben a könyvben Kurzweil többféleképpen reagál. Először is kiemeli jóslatainak sikereit: például az 1990-es években azt jóslta, hogy egy számítógép 1998-ra legyőzi az embert sakkban (és valójában az IBM Deep Blue 1997-ben legyőzte Kasparovot), és előre jelezte az internethasználat, a mobil számítástechnika és az MI népszerűségének robbanásszerű növekedését – mind lényegében helyes. Emlékezteti az olvasókat, hogy *1999-ben a szakértők szerint az AGI 100 évre van, míg ő 30 évre van* – és most sokan egyetértenek az ő időkeretével. Ez egy kicsit olyan, mint "én mondtam." Másodszor, azokon a területeken, ahol a dolgok lemaradtak (gyakran társadalmi/intézményi tényezők miatt), Kurzweil azt állítja, hogy a gyorsító hozamok törvénye még mindig a helyes úton tart, de az örökbefogadásnak lehetnek késései. Ahogy korábban említettük, elismeri, hogy az intézmények lassan változnak, de úgy véli, *hogy végül az exponenciális technológia győz*. Ezt kifejezetten megemlíti, hogy 2005-ös előrejelzése közül néhány olyan eseményt tapasztalt, amit ő *"metaforikus 'ablak bedeszkázásának'"* nevez – vagyis társadalmi ellenállásnak –, majd azt mondja, hogy ezek az akadályok átmenetiek. Például azt mondhatja: *Igen, a teljesen autonóm autók 2024-ben nem mindenütt jelen vannak, de a technológia működik – most a szabályozási és kulturális elterjedés kérdése van, ami közeleg*. A Cassandra párbeszédben Cassandra valószínűleg a túlzott optimizmussal foglalkoztatja őt, és Kurzweil valószínűleg exponenciális trendeket mutató adatokat mutat, amelyek még mindig a következményekre mutatnak, néhány évvel később is néhány évvel később is a következményekre. Fontos megjegyezni, hogy Kurzweil megerősíti, hogy nagy dátumai (2029, 2045) megmaradnak. Ezzel közvetlenül visszautasítja azokat a kritikusokat, akik azt várták, hogy kiszorítsa a Szingularitás dátumát. Úgy véli, hogy az MI-ben most tapasztalható fejlemények bizonyítéka annak, hogy az ütemterv helyes.
- "Figyelmen kívül hagyod a kockázatokat és hátrányokat." Egy másik kritika, hogy Kurzweil annyira rajongott a technológiához, hogy bagatellizálja a kockázatokat (az MI ellenségessé válása, genetikai manipuláció hibái stb.) vagy a társadalmi problémákat (munkahelyvesztés, egyenlőtlenség). *A Szingularitás közelebb* című művében Kurzweil nagyon egyértelműen megpróbálja megmutatni, hogy *nem hagyja figyelmen kívül ezeket*. A 7. fejezetet a "Veszélynek" szenteli, és a legrosszabb forgatókönyvekkel foglalkozik (ahogy fentebb összefoglalták). A Guardian interjújában elismeri az AI etikai

problémáit, mint az elfogultság, mélyhamisítás és visszaélés, megjegyezve: "*Közeleg a választás, és a 'deepfake' videók aggodalomra adnak*", stb.. Ezzel ezzel felhossa a figyelmét, hogy tisztában van azoknak az aggodalmokkal, amelyeket sok kritikus (például Gary Marcus, a néhai Stephen Hawking vagy mások) felvetett. Kurzweil válasza a mérséklést hangsúlyozza a tiltás helyett. Gyakran állítja, hogy *figyelni kell az MI-t és biztosítani kell a fejlődést, de nem kell megállítani a haladást*. A közös erőfeszítéseket (például a szabályok készítésének részeként) említi, hogy bebizonyítja, hogy részt vesz a biztonsági közösségben, nem pedig teljesen tudatlan. A könyv utolsó párbeszédében Cassandra karaktere valószínűleg megtestesíti azok nézetét, akik félnek az irányíthatatlan technológiától, lehetővé téve Kurzweil, hogy ezt közvetlenül megcáfolja azzal, hogy elmagyarázza, *hogyan* tudjuk irányítani, és miért nem valószínű az AI apokalipszis szerint, ha bölcsen kezeljük a dolgokat. Egyenesen szembeszáll a luddita érvelet: leírja az eredeti ludditákat és azt, hogy a technológia végül hogyan szolgált a társadalomnak, ezzel arra utalva, hogy a mai technológiai szkeptikusok ("fundamentalista humanisták") is túl pesszimistáknak bizonyulnak. Lényegében Kurzweil tiszteletben tartja a *kritikusok által felvetett kérdéseket*, de ellentétes következtetésekre jut a *válaszokkal* kapcsolatban. Azáltal, hogy ezeket a vitákat a könyvbe illesztette, azt mutatja, hogy átgondolta a kritikákat, és készen áll a válaszok.

- "Ez olyan, mint egy vallás vagy fantázia." Néhány ellenző szerint a Szingularitás fogalma inkább misztikus elragadtatás, mint tudomány (utalva arra, ahogyan Kurzweil a transzcendenciáról beszél, stb.). Válaszul Kurzweil *lehetőség szerint adatokra alapozza érveit*. A könyv tele van diagramokkal, lábjegyzetekkel és példákkal, amelyek minden állítást alátámasztanak. Például, amikor azt mondja, hogy az intelligencia milliószorosára bővül, ezt összekapcsolja a számítási és agyszkenelési felbontás tényleges trendjeivel. Lényegében azt mondja: *ez nem hit, hanem ismert tudomány extrapolációja*. A metaforákat is tisztázza: például elmagyarázza, hogy a "Szingularitás" kifejezés egy fizikából származó metafora, ahol a normális szabályok meghibásodnak, nem természetfeletti esemény. Konkrét forgatókönyveket kínál (például hogyan működnének a nanobotok, hogyan tudna MI eltávolítani a betegségeket stb.), és megpróbálja eloszlatni azt a nézetet, hogy ez mind vad fantázia. Még a "sci-fi hangzásnak hangzik" kritikát is megválaszolja, hogy sok korábbi sci-fi hangzású jóslat (például valaki 1900-ban azt mondta, hogy az ember a Holdon jár) a technológia révén valóra vált. Ha valami, Kurzweil hangvétele ebben a könyvben megfontolt – lelkes, de elemző –, hogy megmutassa, ez komoly előrelátás, nem pedig technológiai kultusz jóslat.
- "Az emberek elveszítik emberi mivoltukat vagy céljukat." Néhány kritikus (különösen azok, akik humanista vagy spirituális irányultsággal rendelkeznek) attól tartják, hogy a gépekhez való egyesülés vagy az örökké élés elveszheti az élet értelmét, vagy hogy Kurzweil víziója az embereket csak információvá redukálja. A könyvben a 3. fejezet és a Cassandra párbeszéd feldolgozza a "Kik vagyunk, amikor részben gépek vagyunk?" kérdés. Kurzweil azt állítja, hogy *mindig is használtunk eszközöket* – a tüztől az okostelefonokig –, és mindegyik a mi kiterjesztésünké vált. Az MI csak egy intimebb eszköz. Híresen azt mondja: "*A technológia az ember része*", emlékeztetve arra, hogy a ruhaviselés, a nyelvi írás stb. "természetellenes" technológiák, amelyeket mi vettünk át, és még mindig emberek vagyunk. A Guardian interjúbán ezt a példát hozta fel: az emberek eredetileg azt mondták: "Nem akarnék mindig telefont magammal", de most a legtöbben igen – és még mindig teljesen embernek tekintjük magunkat. Tehát a augmentációt természetes evolúcióként értelmezi, nem pedig dehumanizációként. A célt illetően Kurzweil válasza az, hogy a cél és a jelentés az, amivé mi vagyunk teremtve. Megszabadulva a túlélési igényektől, az emberek magasabb célokra (egy

nagyon transzhumanista nézet) összpontosítanak – az univerzum felfedezésére, művészet létrehozására, tudás bővítésére. Megemlíti, hogy még ma is sokan találnak célt az alapvető munkán túl, ha lehetőségük van rájuk. Brett Hurt kommentárján keresztül olyan gondolkodókat idézve meg, mint Viktor Frankl, hangsúlyozza, hogy a jelentés kulcsfontosságú, de nem szükséges ahhoz, hogy hagyományos nehézségek létezzenek. Értelmet fogunk találni a bővített lehetőségeinkben. Ezen felül Kurzweil a "Istent játszom" kritikát is megemlíti (amit túlzásba esünk). A könyvben és az interjúkban gyakran válaszol, hogy a technológia természetes az emberek számára – hogy az univerzumban az intelligencia növelése valójában nemes törekvés, talán akár sors is. Megemlítheti (ahogy a beszélgetéseiben is), hogy *az evolúció adott nekünk egy intelligencia szikráját, és mi ezt használjuk az intelligencia felerősítésére – ez nem kivétel, hanem az evolúció pályájának folytatása.* Ezért azt, amit egyesek "Istenjátéknak" nevezik, *ő intelligens lényként való felelősségünk teljesítésének* nevezi, hogy javítsuk az életet.

- "A korábbi jóslateid csak szerencsések voltak, vagy te válogattál." A kritikusok néha azt állítják, hogy Kurzweil kiemeli a sikereket, és figyelmen kívül hagyja a hibákat. *A The Singularity Is Nearer* című részben néhány korábbi jóslatát is tárgyalja. Például *"értékeli 1999-es előrejelzését"* a 2029-ről a jelenlegi előrelépések fényében (az Amazon összefoglaló ezt valójában megjegyzi). Nyíltan megjegyzi, mely történetek korábban (például az MI érdeklődés), és mely kicsit lemaradtak az időbeosztásban. Egy megjelent könyvben ezzel valójában felelősségre vonja magát (a legtöbb futurista kerüli a lejárt jóslatainak kiemelését!). Kurzweil érvelése szerint az általános trend a jóslat módokon, még ha néhány konkrét részlet is eltért. Valószínűleg említi, hogy az AI 2029-es dátuma jóval visszaesett, és pontosnak tűnik, ami hitelességet ad a 2045-ös dátumához. Megemlítheti azt is, hogy a *The Singularity Is Near* (2005) című könyvből származó 2020-as előrejelzések valóban valóra váltak: például előre jelezte a viselhető számítógépek, az állandó nagy sáv szélességű hálózatok, az MI asszisztensek megjelenését – mind ma már általánosak. Ezekre emlékeztetve ellensúlyozza azt a narratívát, hogy leginkább tévedett.

Összefoglalva, Kurzweil a *The Singularity Is Nearer* című könyvet használja arra, hogy a kritikusokat saját feltételeik szerint megszólítsa: a szövegben felveti a nehéz kérdéseket (strukturált érvek és fiktív párbeszéd révén), adatokat és logikus érvelést nyújt ezekre a válaszokhoz, valamint elismeri az érzelmi vagy erkölcsi aggályokat, miközben azt állítja, hogy azok megoldhatók. Az eredmény Kurzweilről van egy kép, mint még mindig szégyentelen optimista, de hallotta a *szkeptikusokat*, és úgy hiszi, kielégítő válaszai vannak. Ahogy egy kritikus megjegyezte, a könyv *"a leghasznosabb forrása azoknak az innovációknak a bőséges, amelyeket az MI lehetővé tehet... A hátrányokat és veszélyeket is kiemelik, de Kurzweil a szívében evangélista – ami vagy radikális reményt, vagy mély szkepticizmust vált ki, attól függően, hogy mi a hajlamod van."* Más szóval, Kurzweil válaszol a kritikákra, de az, hogy meggyőzőnek találjuk-e a cáfolatait, gyakran a saját technológiai nézetünkön múlik.

Fogadtatás és kritikák

A *The Singularity Is Nearer 2024 közepén megjelent*, és számos reakciót váltott ki szakértőktől, médiumoktól és közvélemény-értelmiségiectől. A fogadtatás vegyes volt, ami tükrözi Kurzweil nézeteinek megosztó természetét:

- Dicséret a vízióért és belátásért: Kurzweil rajongói dicséretet adtak a könyvnek a jövő technológiájának átfogó áttekintéséért és reményteljes üzenetéért. Például Brett A. Hurt

vállalkozó és író "*mély új könyvnek*" nevezte, sőt "*intellektuális túlélési útmutatónak*" is nevezte a közelgő átalakulások megértéséhez. Az ilyen olvasók Kurzweil technológiai trendekből és pozitív előrejelzésekből álló összeállítását inspirálónak találják. Úgy érvelnek, hogy a könyv átfogó bevezető az MI-hez és az exponenciális technológiákhoz, hasznos bárki számára, akit érdekel a technológia jövője. A technológiai vezetők gyakran tisztelik Kurzweil múltját: Bill Gates azt mondta: "*Ray Kurzweil a legjobb ember, akit ismerek az AI jövőjének előrejelzésében*", amit gyakran idéznek Kurzweil munkájáról szóló vitákban. Ez az ajánlás arra utal, hogy a technológiai közösségben Kurzweil előrejelzéseit komolyan veszik. Néhány kritika azt is értékelte, hogy a *The Singularity Is Nearer* frissíti Kurzweil korábbi munkáit aktuális példákkal, így összetett témákat is hozzáférhetővé tesz. Az olyan dolgokról adott magyarázatainak világossága, mint az MI, agykémia vagy nanotechnológia, pozitív jegyeket hozott; még szkeptikus bírálók is néha elismerték, hogy Kurzweil "*összetett témák akadálymentes magyarázatait nyújt*".

- Kritika az új ötletek hiánya miatt: Másrészt több kritikus és kommentátor úgy érezte, hogy a könyv nem kínál sok újat Kurzweil korábbi írásain túl. Megjegyezték, hogy a folytatás lényegében újraértelmezi a 2005-ös könyv érveit frissített adatokkal. Ahogy egy összefoglaló fogalmazott, Kurzweil "*a legújabb technológiai fejlesztésekkel erősíti [a régi ötleteket]*", de alapvetően nem mutat be új keretrendszereket. Ez arra készítette, hogy néhányan megkérdőjelezzék a könyv értékét azok számára, akik már ismerik Kurzweil munkásságát. Ha 2005-ben olvastad a *The Singularity Is Nearer*-t, a *Nearer fő üzenete* nagyon ismerős lesz. Ezt elszalasztott lehetőségként értelmezték, hogy mélyebben foglalkozzunk néhány filozófiai kérdéssel, vagy szükség esetén módosítsák az idővonalat. A kritikusok ebben a táborban gyakran azt akarták, hogy Kurzweil jobban foglalkozzon azzal, miért nem történtek meg bizonyos dolgok (például teljes MI-vezérelt forradalmak) még 2024-ig, vagy hogy több külső nézőpontot is beépítsen. Ehelyett azt kapták, amit szerintük "*Kurzweil legnagyobb slágerei, frissítve.*" Azoknak az olvasóknak, akik új betekintésekre vágnak, ez csalódást keltő volt.
- Túl utópikus / Problémák elsiklása: Sok szkeptikus továbbra sem győz meg Kurzweil optimizmusában. Óvatosabb szakértők véleményei rámutatnak, hogy Kurzweil "*elsikli a lehetséges kockázatokat*" és a társadalmi kihívásokat. Például Gary Marcus, egy MI-kutató, aki kritikusan a szingularitás eképével szemben, azt állíthatja, hogy Kurzweil alábecsüli a tudatosság összetettségét vagy túlbecsüli a fejlődés ütemét olyan területeken, mint az autonóm robotika. Néhány kritikus ismételteti a Newcity kritika által kihangzott pontot: hogy Kurzweil elismeri az akadályokat, de gyakran túl lazán integeti el azokat. Az aggodalom az, hogy nem foglalkozik teljesen azzal, milyen rendezetlenek és makacsok lehetnek az emberi intézmények és viselkedések. Ahogy az a kritika is megjegyezte: "*Hiányzik az emberi természet és az emberi társadalom megértése itt.*" Olyan példákat hoztak fel, mint azt feltételezni, hogy az emberek gyorsan elfogadják a nanobotokat a testükbe vagy a laboratóriumban termesztett húst a tányérjukon, egyszerűen azért, mert ez logikus – míg a valóságban a kulturális és érzelmi ellenállás hosszú ideig tart. Ezek a kritikusok úgy vélik, Kurzweil jövője technikailag megvalósítható, de *szociológiailag naiv*.
- Média és közéleti értelmiségiek részvétele: A fősodratú média Kurzweil könyvét és ötleteit vonzalmas és szkepticizmus keverékével foglalkoztatta. A Guardian/Observer egy nagy visszhangot adott Kurzweil-lel a könyv megjelenése alkalmából, címmel "*2045-ig milliószeresére fogjuk bővíteni a hírszerzést*". A cikk megjegyezte, hogy "néhány előrejelzése már nem tűnik olyan furcsának" a mesterséges intelligencia közelmúltbeli fejlődése miatt, ami azt az érzést tükrözi, hogy Kurzweil ötletei kissé a főáram felé mozdultak. De ugyanakkor rákérdezett, miért bízhatnak az embereknek a

randitársaiban, és hogyan kezeljék a kockázatokat, ami tükrözte a maradék szkepticizmust. A Science Friday, egy népszerű tudományos rádióműsor, 2024 júniusában mutatta be Kurzweilt, így platformot adott számára, hogy széles közönségnek elmagyarázza a könyv témáit. A hangnem tiszteletteljes kíváncsiság tükröződött – elismerve Kurzweil futurista státuszát, miközben felvetette azt a kérdést, hogy "nem ez egy kicsit 'mondtam'?" kérdést az AI felemelkedése fényében. A Wired magazin interjút készített Kurzweil-lel is (a Science Friday oldal, amely egy Wired interjúhoz hivatkozott), valószínűleg technikai pontokat és kételyeket vizsgálva. A Popular Mechanics, egy jövőbeli technológiai lelkesedéstről ismert magazin pedig egy cikket közölt, amelyben Kurzweil kulcsfontosságú állításait – 2045-ig való egyesülés, intelligencianövelés stb. – érdekes, bár óvatos hangnemben mutatta ki.

- A közvélemény: A technológiailag jártas közönség körében (például a Reddit vagy a Twitter fórumokon) a könyv vitát váltott ki. Néhány olvasó az r/singularity-n és hasonló fórumokon izgatott volt, hogy Kurzweil frissíti előrejelzéseit, és élvezte a fejezetről fejezetre szóló beszélgetéseket (a Reddit poszt, amely összefoglalta az 1. fejezetet, olyan emberek körében is érdeklődött, akik az agy-számítógép interfészek megvalósíthatóságáról beszéltek). Mások csalódást keltettek, kijelentették, hogy "valójában nincs semmi új, csak az utolsó fejezet", és hogy Kurzweil "érvelt az emberi elme relevanciájáról", de más részeket kihagy (ezt a gondolatot Carlos Perez mesterséges intelligencia kutató Twitter-számban látta). A Goodreads-en a könyv átlagos értékelése körülbelül 3,9/5-ös, ami elég jó fogadtatást jelez, de nem kiemelkedő. Sok 4- és 5 csillagos értékelés dicséri, mint az elmét bővítő, míg az alacsonyabb értékelések gyakran ismétlést vagy túlzott optimizmust említenek indokként. Az egyik oldal összefoglalója így szólt: "A vélemények megosztottak... sokan Kurzweil jóslatait lenyűgözőnek találják, de [néhányan] úgy érzik, hogy hiányzik a társadalmi következmények és etikai kérdések mélysége a kezelésében".

Lényegében a *The Singularity Is Nearer* gondolatébresztő, de megosztó könyvként fogadták. A rajongók a jövő víziós tervének nevezik (a Wall Street Journal nem értékelte hivatalosan, amennyire tudom, de olyan személyiségek, mint Peter Diamandis vagy Tesla közössége valószínűleg pozitívan beszéltek). A kritikusok techno-utópikus ösztönzőnek nevezik, amely nem küzd a valósággal. És sok olvasó a középső helyen áll, értékelve Kurzweil képzeletét és enciklopédikus technológiai trendeket ismerősségét, de jóslatait fenntartással kezeli. Ahogy egy kritikai összefoglaló találóan fogalmazott, Kurzweil olvasása "radikális reményt vagy mély szkepticizmust" válthat ki, attól függően, hogy milyen hajlamot érzel. A könyv fogadtatása így tükrözi az MI-ről és a jövőről szóló szélesebb diskurzust – az izgalom és az aggodalom keverékét.

Figyelemre méltó interjúk és előadások a könyv népszerűsítését

A *The Singularity Is Nearer* népszerűsítéséért Ray Kurzweil 2024 folyamán több interjút és nyilvános előadást tartott, amelyben a könyv tartalmáról és előrejelzéseiről beszélt:

- SXSW 2024 (Austin, 2024 március): Kurzweil 2024 márciusában tartott kiemelt előadást a South by Southwest konferencián, néhány hónappal a könyv megjelenése előtt. Egy beszélgetésben Nick Thompsonnal (az *Atlantic* vezérigazgatója) Kurzweil előzetesítette a könyv számos témáját. Ez az előadás nagy felhajtást keltett; *SiliconHills News* egy cikket közölt "7 kulcsfontosságú tanulság Ray Kurzweil SXSW-n tartott előadásából", amely összefoglalta fő pontjait. Néhány tanulság:

1. Szingularitás 2045-re: Megismételte, hogy az MI felülmúlja az emberi intelligenciát, és 2045 körül egyesülünk vele.
2. Moore-törvény és MI fejlődés: Az exponenciális számítástechnikai növekedés olyan áttöréseket hajt végre, mint az LLM-ek, és az AGI-hez vezet.
3. Optimizmus az MI előnyeiről: Hangsúlyozta, hogy az MI nagyban fog segíteni olyan területeken, mint a hosszú élet, a szegénység és betegségek megszüntetése, valamint az általános vagyonteremtés.
4. Agy-számítógép interfészek: Az AGI elérése után olyan technológiák, mint a BCI és a nanotechnológia, elengedhetetlenek ahhoz, hogy az emberek és az MI valóban összeolvadjanak.
5. Etikai kockázatok (Paperclip Maximizer): Elismerte a "paperclip" gondolatkísérletet (egy MI, amely a végtelenségig követi a célt), mint egy olyan dolog, amivel szembe kell szállni, ami bizonyította, hogy tisztában van az AI kockázati helyzeteivel.
6. Tudatosság az MI-ben: Azt állította, hogy *a tudatos élmény nem tudományosan meghatározható akadály*, utalva arra, hogy az agy kapcsolatainak lemásolása elegendő az elme újraterejtéséhez (nincs szükségünk misztikus szikrara).
7. Elmementés és digitális halhatatlanság: A Szingularitás után az emberek képesek lesznek visszamenteni az elméjüket, és gyakorlatilag örökké élni fognak azzal, hogy szükség esetén visszaállítják a mentéseket.

A SXSW felvételtől egy videót is feltettek a YouTube-on ("The Singularity Is Nearer featuring Ray Kurzweil | SXSW 2024"), amely bővíti elérését. Az SXSW-n való megjelenés azért volt jelentős, mert Kurzweil ötleteit nagy technológiai közönség és média elé tette, ami sok vitát és visszhangot váltott ki jóval a könyv megjelenése előtt.

- TED 2024 (2024. április, Vancouver): Ray Kurzweil szintén beszélt a TED 2024 konferencián. A TED weboldalán Ray Kurzweil *"Az elmúlt 6 évtized mesterséges intelligencia — és mi következik"* című előadást is felvet, amely feltehetően szorosan kapcsolódik a könyv tartalmához. A TED-en Kurzweil elmesélte az AI előrehaladását (60+ éve van benne), majd valószínűleg összefoglalta a következő 20 évre vonatkozó előrejelzéseit (ami 2045-ig vezet). Ez fontos platform volt, mert a TED előadásokat széles körben nézik; Segített a könyv ötleteinek globális népszerűsítésében. A TED oldalán szerkesztői megjegyzés szerint *a The Singularity Is Nearer* hamarosan megjelenik, Kurzweil előadását közvetlenül a könyvhöz kötöttnek helyezve. Ebben az előadásban Kurzweil valószínűleg sok pontot visszhangzott a könyvből – beleértve a számítási fejlődést (gyakran bemutatja híres dolláronkénti számítási grafikonját évtizedek alatt), a hat korszakot és a jövőképét. A TED nem adott ki teljes átiratot a cikk írásának idejére, de ismert, hogy Kurzweil TED-fellépései befolyásosak (korábbi TED-előadásai milliós megtekintéseket gyűjtöttek).
- Médiainterjúk (nyomtatott/online): Kurzweil különböző médiumokkal foglalkozott:
 - The Guardian/Observer interjú (2024. június 29.): Ez a korábban idézett részletes kérdezz-felelek során Kurzweil válaszolt arra a kérdésre, hogy miért írta a folytatást, és összefoglalta a kulcsfontosságú pontokat: az MI elérése 2029-re még mindig elérhető, 2045-re a felhővel való egyesülés stb.. Beszélt a jelenlegi AI problémákról is, mint az LLM hallucinációk, és arról, hogyan fogják ezeket megoldani. Az interjúztató Zoe Corbyn megkérdőjelezte őt a randiban való hittelt, és ő megmagyarázta az indoklását. Beszéltek továbbá a munkahelyi hatásokról (ahol a 2030-as években hozta fel az UBI-t), és az élettartamhosszabbítást (ahol említette az orvosi nanobotokat a 2030-as

években, és az elme-feltöltést a 2040-es években). Ezt az interjút széles körben olvasták, újra közölték vagy más médiumokban idézték, hozzájárulva a könyv nyilvános diskurzusához.

- Science Friday (NPR Radio, 2024. június 14.): Kurzweil szerepelt a *Science Friday*-ben Ira Flatow-val (és Annie Minoff társműsorvezetővel mutatta be őt). Ebben az interjúban (hanganyag és átirat elérhető ScienceFriday.com-n keresztül) Kurzweil arról beszélt, hogyan változott a világ 2005 óta, és miért volt szükség a folytatásra. A műsorvezetők játékos, de éles kérdéseket tettek fel, mint például: "Nehéz nem csak annyit mondani, hogy 'Mondtam?'" az AI pályájára utalva, így Kurzweil elmesélhette, hogyan alakultak előrejelzései. Elmagyarázta a számítástechnika exponenciális gráfonját, és azt, hogy ez hogyan áll előrelátása alapja. Valószínűleg személyes elemeket is érintett (Kurzweil gyakran említi, hogy étrend-kiegészítőket szed, vagy egyszer mesterséges intelligencián keresztül akarja feltámasztani az apját – bár nem vagyok benne biztos, hogy ez az interjú erre ment-e). Az NPR Science Friday műsorában való részvétel széles platformot adott számára a tudomány iránt érdeklődő általános közönség számára.
- Wired interjú (2024 körül): Ahogy a Science Friday linkje is mutatja, a Wired adott interjút vagy cikket Kurzweilről. Bár nincs meg a szövegem, a Wired általában okos kérdéseket tesz fel, talán az AI rövid távú problémáiról és Kurzweil nézeteiről a legutóbbi AI fejlesztésekről (ChatGPT stb.). A Wired történelmileg keveredett a csodálattal és szkepticizmussal szemben Kurzweil iránt, így ez egy érdekes cikk lett volna.
- Egyéb médiumok: Olyan médiumok, mint a CNBC vagy a Bloomberg, talán interjút készítettek vele, különös tekintettel az iparágra gyakorolt következményekre, bár én nem láttam részleteket. A The New York Times *Book Review* is röviden hivatkozott a könyvre egy MI könyvekről szóló trendcikkben (tekintve, hogy 2023–24-ben MI könyvek árasztották el). És olyan technológiai oldalak, mint a Singularity Hub (amely a Singularity University-hez kötődik, Kurzweil társalapítója) kétségtelenül pozitív értékelést adtak a könyvnek.
- Podcast-megjelenések: Kurzweil feltűnhetett néhány tech podcastban vagy YouTube műsorban. Például Lex Fridman, aki gyakran interjút készít MI gondolkodókkal, valószínű helyszín lehet (bár nem emlékszem konkrét epizódra Kurzweil-lel 2024-ben; Lex 2019-ben interjút készített Kurzweil-lel). Ha podcastokat csinálna, azok hosszabb formátumú beszélgetést engedtek volna a könyvről. Hasonlóképpen, a "Honestly" podcastban Bari Weiss-szel Steven Pinker az MI optimizmusról és a pesszimizmusról beszélt (Kurzweil ötleteit ebben a kontextusban említették Brett Hurt cikke szerint); lehetséges, hogy maga Kurzweil is megjelent egy ilyen fórumon, hogy megfogalmazza álláspontját az AI végzet narratívájával szemben.
- Könyvelőadások és panelbeszélgetések: A megjelenés után Kurzweil valószínűleg virtuális vagy személyes könyvelőadásokat tartott. Például beszélhetett a Commonwealth Clubban vagy egy New York-i Közkönyvtári rendezvényen. Egy esemény a feljegyzés: a Singularity University (Kurzweil és Peter Diamandis alapítója) valószínűleg könyvbemutató webináriumot vagy beszélgetést szervezett. Emellett Kurzweil tartott egy beszélgetési eseményt David S. Rose-szal (egy angyalbefektető), ami YouTube-on van, ami lehet, hogy az SXSW egyik vagy másik eseménye. Ezen felül az EDRM blogbejegyzés, amit láttunk, Ralph Losey volt, aki említette, hogy Kurzweil 2024. június 28-án beszélt, és még utalt egy Kurzweil hangjának videós avatárjára, amelyet ő készített. Ez arra utal, hogy Kurzweil esetleg adott időpontban (június 28.)

virtuális előadást vagy interjút tartott – ami lehetett a TED előadás vagy egy privát szekció –, és a rajongók részt vettek benne.

- Nyilvános kérdezz-felelek és közösségi média: Kurzweil maga nem túl aktív a közösségi médiában, de a könyvről szóló beszélgetések felbukkantak a Twitter/X-en (például @IntuitMachine (Carlos Perez) szála, amit láttunk, és olyan gondolkodók posztjai, mint Max Tegmark vagy Sam Altman, amelyek talán Kurzweil idővonalára utalnak kommentárokból). Sam Altman (az OpenAI vezérigazgatója) korábban megfigyelte Kurzweil előrejelzéseit; a ChatGPT sikere miatt Altmant megkérdezték, hogy rövidebb-e az AGI idővonala, amire óvatosan válaszolt, de Kurzweil 2029-es dátuma gyakran szerepel ezekben a vitákban. Figyelemre méltó, hogy Kurzweil ötletei közvetve részei voltak az *MI szünetének* nyílt levél vitájának 2023 elején – sokan, akik aláírták azt a levelet (aggódtak a gyors MI fejlődés miatt), hallgatónan azt mondták: "Kurzweil szingularitása gyorsabban érkezik, mint vártuk, és félünk." Kurzweil maga a könyv népszerűsítésével álláspontot foglalt az AI szünete ellen, hangsúlyozva a megfigyelést, nem pedig a moratóriumot. Valószínűleg ezt megemlítette a beszélgetéseken, összhangban azzal a kérdéssel, mint "*folytatnunk kell az MI fejlesztését a hasznák érdekében*".

Összességében Kurzweil 2024-ben a könyv bemutatásakor különösen *a tavasszal és kora nyaron jól látható volt*. Az SXSW-től a TED-ig és a nagy médiainterjúkig következetesen közvetítette a könyv kulcsüzeneteit mind a technológiai bennfenteseknek, mind a nagyközönségnek. Minden megjelenés megerősítette a kiemelkedő futurista imázsát, aki bár akkoriban a hetvenes évei közepén járt, még mindig lelkesen tekint a következő évtizedekre. Ha más nem, a *The Singularity Is Nearer* promóciós körköre megmutatta, hogy Kurzweil és gondolatai továbbra is rendkívül befolyásosak és lenyűgöző témák, ahol beszél, vitákat váltanak ki.

Források:

- Pfeiffer, Scott. "Közös barátunk: A szingularitás közelebb kerül értékelése: Amikor egyesülünk az AI-val, Ray Kurzweil." *Newcity Lit*, 2024. október 14.
- Soffer, Alon (összefoglaló). *A szingularitás közelebb van* – összefoglaló és GYIK. *SoBrief* (frissítve: 2025. február 17.).
- Corbyn, Zoë. "Ray Kurzweil mesterséges intelligencia tudós: '2045-re milliószeresére fogjuk bővíteni az intelligenciát'." *The Guardian (Observer)*, 2024. június 29.
- Hurt, Brett A. "Az MI-vezérelt univerzum egy szempillantásra van." *Medium*, 2024. július 11.
- Losey, Ralph. "Ray Kurzweil új könyve: *A szingularitás közelebb van* (amikor egyesülünk az AI-val)." *EDRM Blog*, 2024. július 16.
- Lorek, Laura. "7 kulcsfontosságú tanulság Ray Kurzweil SXSW-n tartott előadásából: *a szingularitás közelebb van*." *SiliconHills News*, 2024. március 29.
- Science Friday szegmens és részlet: "*A szingularitás közelebb, 'mondja Ray Kurzweil futurista*." *Science péntek*, 2024. június 14.
- Reddit beszélgetés: "A szingularitás közelebb van, Ray Kurzweil, 1. fejezet – Hol vagyunk?" *r/singularity*, 2024

Kérjük, részletesen ismertesse Yuval Harari új könyvét, a "Nexus"-t.

Yuval Noah Harari *Nexus*: Átfogó áttekintés

Nexus összefoglalója és szerkezete

A *Nexus: A Brief History of Information Networks from the Stone Age to MI* Yuval Noah Harari legújabb ismeretterjesztő műve, amely 2024 végén jelent meg. A könyv ~528 oldalt foglal fel, és nyomon követi, hogyan formáltak az információs hálózatok az emberi történelmet – az őskori törzsektől a mesterséges intelligencia korszakáig. A könyv egy előszóba, 11 fejezetbe és egy epilógusba épül, amely átfogó történelmi narratívát ötvözi a kortárs elemzéssel. Minden fejezet egy kulcsfontosságú szakaszt vagy aspektust vizsgál az információáramlás társadalmi részéről, megerősítve Harari érvét, miszerint az információ kontrollálása *megteremtette és tönkretette a világunkat*. Az alábbiakban fejezetről fejezetre olvasható a tartalom és a szerkezet:

- 1. fejezet – *Mi az információ?* – Bemutatja az információ fogalmát és annak szerepét az emberi ügyekben. Harari megkérdőjelezi azt a "naiv nézetet", miszerint több információ szükségszerűen több igazsághoz vagy bölcsességhez vezet. Azt állítja, hogy az információ elsődleges funkciója, hogy az embereket hálózatokba kötse, nem pedig az igazság feltárása – valójában az emberek gyakran fikció révén kötnek össze a kötélezve, nem pedig a tényeket. Ez a fejezet felállítja a központi paradoxont: az emberek a legtájékozottabb fajok, mégis a rossz információk ismételt önpusztító viselkedésbe vezetnek minket.
- 2. fejezet – *Történetek*: A közös narratívák és mítoszok erejét vizsgálja az emberi együttműködés bővítésében. Harari a *Sapiens* témájára épít, miszerint a fiktív történetek (vallások, ideológiák stb.) nagyszabású együttműködést tesznek lehetővé a kis rokonsoportokon túl. Elmeséli, hogyan válogatták és szentelték a szent szövegeket, mint a Biblia, – ez egy korai információkontroll formája, amely közös hiedelmek alatt egyesítette a társadalmakat. Harari azzal, hogy az őskori szóbeli hagyományok és későbbi vallási mítoszok hogyan kötöttek össze a közösségeket, Harari azt mutatja, hogy a kollektív *hiedelmek* alkotják az emberi információs hálózatok gerincét.
- 3. fejezet – *Dokumentumok*: Az írás, a nyilvántartás és a bürokrácia megjelenését tárgyalja. Harari azt állítja, hogy az írásos dokumentumok lehetővé tették a központosított államok és birodalmak felemelkedését azáltal, hogy lehetővé tették az információk (például adónyilvántartások vagy törvények) tárolását és továbbítását időn és távolságon keresztül. Korai példák (sumér gabonatóblák, népszámlálási listák, jogi kódexek) megmutatják, hogyan vált az adatgyűjtés és a papírmunka a kormányzás

eszközeivé. Harari kiemeli a sötét oldalt is: akik a dokumentumokat irányítják, az emberek életét is irányíthatja. Megható példát hozol a fasiszta korszak Romániájának (1930-as évek), ahol a zsidóknak (köztük Harari nagyapjának is) állampolgársági papírokat kellett bemutatniuk, amelyeket a hatóságok szándékosan megsemmisítettek – ez egy bürokratikus trükk volt, amely sokakat gyakorlatilag állampolgárság nélkül hagyott. Ez jól szemlélteti, hogyan *lehet a papírmunkát és archívumot* fegyverként használni hatalmi információs hálózatokká.

- 4. fejezet – Hibák: Vizsgálja, hogyan terjednek a félretájékoztatás, hibák és hazugságok az információs hálózatokon keresztül. Harari megkérdőjelezi azt a feltételezést, hogy az új média mindig előmozdítja az igazságot, és a nyomdát esettanulmányként használja. Miközben a sajtó segítette a tudás terjesztésében, ugyanígy tette lehetővé a *hamis hírek* terjedését a korai modern Európában – például a nyomtatott boszorkányüldöző kézikönyvek és szórólapok tömeges hisztériát terjesztettek a boszorkányokról, és szították a borzalmas boszorkányüldözéseket. Harari lényege, hogy minden információs forradalom (még a látszólag pozitív is) kockázattal jár a hibák és téveszmék terjesztésének kockázatával. A fejezet latin epigráfja ("Tévedni emberi, a hibában kitartás ördögi") hangsúlyozza, milyen kitartóak lehetnek a hamis hiedelmek, különösen, ha új médium erősíti őket.
- 5. fejezet – *Döntések*: Arra fókuszál, hogyan befolyásolja az információáramlás a döntéshozatalt a társadalmakban. Harari összehasonlítja az autoriter és demokratikus információs hálózatokat, és azok döntéshozatalát. Diktatúrákban az információt szigorúan ellenőrzik – gyakran rossz döntésekhez vezetnek, mert a vezetők szűrt vagy hamis adatokat kapnak. (Például utal történelmi esetekre, amikor autoriter rezsimek kényelmetlen igazságokat rejtettek el, ami katasztrófához vezetett.) Nyitottabb rendszerekben a döntések szélesebb információs bemenetekből profitálnak, de félretájékoztatás megbéníthatja őket. Ez a fejezet történelmi és modern példákat használ arra, hogy megmutassa: az információs csatornák kialakítása – ki tudja mit – közvetlenül befolyásolja a döntések minőségét a legmagasabb szinten.
- 6. fejezet – *Az új tagok*: A gépek és az MI információs hálózatok résztvevői megjelenését örökíti. Harari a 20. század közepén megjelenő számítógépek megjelenését és olyan úttörők ötleteit, mint Alan Turing, meséli el. Leírja, hogy a történelemben először az *emberi ügynökök* (algoritmusok, számítógépek) elkezdtek információt feldolgozni, sőt döntéseket is hozni. Ez a fejezet a korai MI kísérleteket és a gépek "intelligens" entitásként való tekintésének fogalmi ugrását tárgyalja. Harari ezeket a technológiákat *hálózataink új tagjaiként* állítja be – a kódfeltörő gépektől a kortárs algoritmusokig –, ami előrevetíti az MI átalakító hatását az emberi együttműködésre és kommunikációra.
- 7. fejezet – *Könyörtelen*: Az autoriter információs rendszerek kitartó természetét vizsgálja. Harari azt vizsgálja, hogyan gyűjtenek adatokat az elnyomó rezsimek és hogyan érvényesítik az ideológiákat könyörtelenül, mindenütt átható módon. Például totalitárius kormányok alatt (például a hidegháború korszakbeli Kelet-Európában) az állami megfigyelés és propaganda végtelen volt, áthatolva az élet minden területére. Harari narratívája valószínűleg magában foglalja a román fasiszta és kommunista rezsimeket is, amelyek a könyörtelen bürokratikus ellenőrzést – például minden lépéshez dokumentumokat vagy igazolványokat követelnek – és a könyörtelen propagandakampányokat is bemutatják. A *"Relentless"* cím azt tükrözi, hogy ha az információs hálózatok (például titkosrendőrség, cenzúrahivatalok stb.) megalakulnak, megállás nélkül működnek a hatalom fenntartásáért. Ez a fejezet háborzongató képet fest az információs hálózatokról, amikor teljes irányítás eszközeként használják.

- 8. fejezet – *Hibás*: Hangsúlyozza, hogy minden információs rendszer (még a fejlett vagy "racionális" rendszerek is) hibázható, és kiemeli az önkorrekció fontosságát. Harari szembeállítja azokat a rendszereket, amelyek tévedhetetlennek állítják magukat (dogmatikus vallások, totalitárius ideológiák vagy akár hibátlannak nevezett mesterséges intelligencia algoritmusok) azokkal a rendszerekkel, amelyek elismerik a hibákat és javítják azokat. Valószínűleg arról beszél, hogy a tudomány és a liberális demokrácia hogyan építi be a visszacsatolást és kritikát – amit ő *"önkorrekciós mechanizmusoknak"* nevez –, míg a zárt rendszerek nem. Történelmi példák (például a Szovjetunió kudarcok elismerésének megtagadása, amelyet Alekszandr Solzsenyicin *Gulag-szigetcsoportja* dokumentál) alámutatják, hogy a hibázás tagadása katasztrófához vezet. Az önkorrekció támogatásával Harari előkészíti az alapot annak érvének, hogy az MI korszak túléléséhez olyan intézményekre van szükség, amelyek tanulni tudnak a hibákból, ahelyett, hogy utópikus tökéletességet hajszolnának.
- 9. fejezet – *Demokráciák*: Az információ szerepét vizsgálja a demokratikus társadalmakban. Harari megjegyzi, hogy maga a demokrácia információáramláson épül – nyilvános vita, szabad sajtó, választások –, amelyek csak olyan újítások révén váltak megvalósíthatóvá, mint a tömegnyomtatás, távirók és műsorszolgáltató média. Megjegyzi, hogy minden médiatechnológiai forradalom (az újságoktól a közösségi médiáig) felfordulást okoz a demokrácia működésében. Ebben a fejezetben Harari a modern kihívásokkal foglalkozik: a közösségi hálózatok és algoritmusok megjelenésével, amelyek gyorsan terjeszthetik a félretájékoztatást vagy szélsőséges eszméket. Valódi esetekről beszél, mint például a Facebook hírfolyama, *amely 2016–17-ben erőszakot szított Mianmarban*, vagy a nyugati demokráciákban elterjedt összeesküvés-elméletek terjedése. Harari azt kérdezi, hogy a demokráciák képesek alkalmazkodni "önkorrekció" intézményeihez a vírusos hamisságok és az MI által generált tartalom korszakához. A téma óvatosan optimista: a demokráciák túlélhetnek, ha megtanulják frissíteni a védelmi intézkedéseket, de még mindig nincs döntés.
- 10. fejezet – *Totalitarizmus*: Ez a fejezet a 9. fejezet ellenpontja, amely a totalitárius rezsimek információs hálózatait vizsgálja, múltbeli és jelenbeli egyaránt. Harari a 20. századi példákat is felidézi, mint a náci Németország és a sztálin Szovjetuniója – olyan rezsimeket, amelyek a hatalmat a propagandára, cenzúrára és aprólékos információkontrollra támaszkodtak. Ezután párhuzamokat vont a modern digitális autoriter rendszerrel: például Kína Nagy Tűzfala és az MI által vezérelt megfigyelés, vagy Irán arcfelismerő alkalmazása, amellyel nyilvánosan elfogják a leplezett nőket. Harari azt javasolja, hogy *míg a diktatúrák elnyomják az igazságot a rend fenntartása érdekében, merev hálózataik végső soron nem rendelkeznek korrekciós visszacsatolással*. Különösen azt állítja, hogy még a diktátorok is küzdenek az autonóm MI rendszerekkel, amelyeket nem tudnak teljesen irányítani (például egy cenzúra szerepben lévő mesterséges intelligenciával, amely *maga* is szembeszállhat a diktátor narratívájával). A fejezet egy nyugtalanító gondolatot emel alá: a fejlett információtechnológia megerősítheti a totalitárius irányítást (megfigyelésen és "álhírek" árvízen keresztül), ugyanakkor magukban rejtheti a kiszámíthatatlanság magvait is, amelyek még a diktátorokat is megijeszti.
- 11. fejezet – *A szilíciumfüggöny*: Az utolsó fejezetben Harari a jövőbe tekint, és figyelmeztet az emberiség információs hálózatában fenyegető szakadásra. Ő alkotta meg a *"Szilíciumfüggöny"* kifejezést (amely visszhangozza Churchill vasfüggönyét), hogy leírja azt a helyzetet, amikor a világ digitális ökoszisztémái szétválnak, például Kína és a Nyugat is összeegyeztethetetlen MI és internet rendszereket épít. Egy ilyen megosztottság "véget vetne az egyetlen közös emberi valóság gondolatának", ahogy az adatok és az igazság geopolitikai blokkok által elszigetelődött. Harari apokaliptikus

lehetőségeket is fest itt. Azt javasolja, hogy az *MI-vezérelt "uralkodók"* isteni erőket szerezhetnek – például egy MI új járdémiás vírust alkothat, vagy destabilizáló deepfake narratívákat hozhat létre, és tömeges erőszakot szíthat. Ebben a helyzetben az MI-nek nem lenne szüksége Terminátor-stílusú robotokra; Információs hálózatokon keresztül manipulálhatja az embereket, hogy pusztító célokat érjen el. A *Szilíciumfüggöny* fejezet lényegében Harari felhívása: ha engedjük az ellenőrizetlen mesterséges intelligencia versenyt és az információs hidegháborút, katasztrofális következményeket kockáztatunk a globális civilizáció számára.

- Utóbeszéd: Harari óvatosan reménykedő hangon zárul. Ismétlem, hogy a történelem nem determinisztikus – a technológia hatása attól függ, milyen döntéseket hozunk. Azt állítja, hogy a legrosszabb következmények elkerüléséhez (mesterséges intelligenciától vagy töredezett valóságoktól) az emberiségnek felelősséget kell vállalnia és *ellenőrzést* gyakorolnia az információs hálózatok felett. Gyakorlatilag Harari az algoritmusok és az MI erős szabályozását, valamint az intézmények megerősítését "önkorrekciós mechanizmusokkal" (mint például a tudomány, független média és demokratikus kormányzás), amelyek képesek megkülönböztetni az igazságot a fikciótól. Az epilógus visszakapcsolódik a könyv nyitó kérdéséhez – ha valóban *Homo sapiens* vagyunk, a "bölc ember", akkor bizonyítanunk kell bölcsességünket azzal, hogy proaktívan irányítjuk az információs forradalmot, ahelyett, hogy hagynánk, hogy elárasztson minket. Ez az utolsó üzenet *a Nexus-t a probléma diagnózisából egy tájékozott cselekvésre való kérésre változtatja.*

Főbb érvek és témák

A Nexus több kulcsfontosságú érvet és visszatérő témát szót a fejezetei során:

- Az információs hálózatok minden emberi rendszer alapja: Harari átfogó tézise az, hogy minden jelentős emberi intézmény – legyen az pénz, vallás, nemzetek vagy vállalatok – alapvetően egy információs hálózat. A nagyszabású együttműködést megosztott információk teszik lehetővé (gyakran történetek, jelek vagy adatok formájában), amelyek több ezer vagy millió egyént kötnek össze. Például a valuták azért működnek, mert kollektív hitel van a pénzügyi információkban; a vallások közös szentírások köré egyesítik az embereket; A nemzetállamok bürokráciákra és hírekre támaszkodnak, hogy megkötsék az állampolgárokat. Ez a téma tovább bővíti *Sapiens* elképzelését, miszerint a "képzelt rendek" (mint a mítoszok és törvények) koordinálják a társadalmat, átformálva azt az információáramlás szempontjából.
- Közös fikciók vs. igazság: Egy visszatérő felismerés a mitológia és a valóság közötti feszültség az emberi hálózatokban. Harari hangsúlyozza, hogy bár az információ hordozhatja az igazságot, *nagyobb ereje közös hitek létrehozásában* rejlik. Valójában a legegyszerűbb módja annak, hogy sok embert összekapcsoljanak, a fikció, nem pedig tények. A történelem során a nagy mítoszok (vallási vagy ideológiai) birodalmakat kötöttek össze – néha az igazság rovására. Ezt a kompromisszumot Harari elképzelésében is összefoglalja, miszerint a társadalom egyensúly *a mítosz és a bürokrácia, a bölcsesség (igazság) és a hatalom (rend)* között. A demokráciák az igazságot a szabad információval támogatják (a káosz kockázatával), míg a diktatúrák a rendet a kontrollált információval (az igazság rovására). A könyv folyamatosan visszatér ehhez a témához: az emberek történeteket használnak társadalmi rend megteremtésére, de ezek a történetek ugyanolyan könnyen lehetnek hazugságok vagy téveszmék.

- Hatalom, téveszme és "Az okos emberek ostobasága": Harari provokatív keretkérdése – *"Ha az emberek ilyen okosak, miért vagyunk mi ilyen buták?"* – a Nexus nagy részét képezi. Válasza az, hogy az emberek egyénileg intelligensek, de társadalmakként hajlamosak vagyunk a kollektív téveszmékre, ha rossz információkat kapunk. A jó emberek katasztrófális döntéseket hozhatnak, ha az információik hibásak. Így az emberiség *hatalmas hatalmat* szerez azzal, hogy milliókat hálózatba épít, de ez a hatalom gyakran törékeny mítoszok, propaganda vagy tévedések alapjain alapul. Az egyik kulcsfontosságú érv az, hogy a legtöbb információs hálózatunk szerkezete miatt rossz kimeneteket eredményeznek (pl. visszhangkamrák, propaganda hurkok), és ahogy hálózataink egyre erősebbé válnak, a *katasztrófa lehetősége* is nő. Harari ezt történelmi katasztrófákkal (például boszorkányüldözéssel vagy totalitárius rezsimekkel) szemlélteti meg, és figyelmeztet, hogy a modern hálózatok (közösségi média stb.) hasonló tömeges ostobasági mintákat mutatnak. A Nexus paradoxona az, hogy a tudás exponenciálisan nőtt, mégis a társadalmak a tudatlanság és félretájékoztatás áldozatai – ezt a mintát Harari annak tulajdonítja, hogy szándékosan és véletlenül *terjesztik a kényelmetlen igazságokat a megnyugtató hamisságokat*.
- Az MI forradalom mint a téveszmék "Nexus": A könyv központi fókuszja a mesterséges intelligencia megjelenése, mint az emberek által létrehozott legösszetettebb információs hálózat. Harari az MI-t potenciális *"idegen intelligenciaként"* ábrázolja – valami nem emberi, amely új módon generál tartalmat, kultúrát és döntéseket. Az egyik nagy érve, hogy az MI új "hálózatot" alkothat a tömeges téveszmék számára a 21. században. Más szóval, ha a korábbi korszakokat az ember által alkotott mítoszok uralták, akkor a következő korszak nagy narratíváit algoritmusok generálhatják. Azt javasolja, hogy az MI egész ideológiákat, propagandákat és hamis valóságokat szülhet, amelyek emberfeletti pontossággal kihasználják pszichológiai gyengeségeinket. Harari az MI hálózatokat "kiszámíthatatlannak és bonyolultnak" nevezi, ami mindenben, amit eddig tapasztaltunk. Feltűnő téma, hogy *az MI-nek nincs szüksége tudatosságra ahhoz, hogy hatalmat gyakoroljon* – az információ (amit látunk és hallunk) irányításával az MI mélyrehatóan alakíthatja az emberi hiedelmet és a politikát. Ez az érv Harari korábbi "dataizmus" fogalmára épül (a *Homo Deus*ből), figyelmeztetve, hogy hatalomot adhatunk át az algoritmusoknak. A Nexus című művében ezt a figyelmeztetést erősíti: az MI információmanipulációs képessége befolyásolhatja szabad akaratomat, sőt akár a liberális demokráciát is, ha nem korlátozza.
- Fenyegetések a demokráciára és a megosztott valóságra: Az MI vitájához kapcsolódik Harari aggodalma, hogy a modern információs hálózatok aláássák a demokráciához szükséges közös igazság alapokat. Megjegyzi, hogy a digitális visszhangkamrák és mélyhamisítványok hogyan törhetik szét a közönséget elszigetelt valóságbuborékká. A *"Szilíciumfüggőn"* (egy jövőbeli megosztott internet) fogalma ezt a fenyegetést dramatizálja. Ha a világ egyik fele egy "igazságot" hisz, a másik fele pedig teljesen más narratívát (az összeegyeztethetetlen hálózatok vagy mesterséges intelligencia propagandának köszönhetően), akkor az informált közbeszéd gondolata összeomlik. Harari ezért a *közös ismeretelméleti alapok* megőrzését azonosítja kulcsfontosságú kihívásként. Ez a téma visszhangozza korábbi "posztigazság" témájú munkáját (a *21 Lessons*-ban), de a Nexus konkrét helyzetekre alapozza ezt, mint például kínai és nyugati MI ökoszisztémák, vagy algoritmusokkal erősített QAnon-stílusú összeesküvések. A könyv sürgető kérdése: Tudnak-e a demokratikus társadalmak alkalmazkodni az új információs környezethez anélkül, hogy káoszba vagy autoriter rendszerbe esnének? Harari azt állítja, hogy lehetséges – de csak tudatos erőfeszítéssel a technológia szabályozására és a közvélemény oktatására, ahogy az alábbiakban tárgyaljuk.

- Önkorrekción és kormányzás szükségessége: A gyakran súlyos előrejelzései ellenére a *Nexus* implicit megoldást kínál: megduplázza az önkorrekción intézményeket és az információ bölcs irányítását. Harari gyakran említi, hogy a tudomány működése közben egy olyan információs rendszert példáz, amely saját hibáit korrekciós lektorációval és bizonyítékokkal javítja – egy modellt, amelyet utánózni kell. Hasonlóképpen, a liberális demokráciákban is vannak visszacsatolási mechanizmusok (szabad sajtó, választások, ellenőrzések és egyensúlyok), amelyek képesek irányirányt javítani a politikákban. Ezeket szembeállítja merev rendszerekkel (szektákkal, diktatúrákkal, még szabályozatlan közösségi média algoritmusokkal), amelyekről hiányzik a korrekciós visszajelzés, és így kicsúsznak az irányítás alján. A könyv egyik visszatérő felhívása az, hogy az MI és technológiai cégeket demokratikus felügyelet alá helyezték – lényegében *elszámoltathatóságot* és emberi értékeket injektálva ezekbe az új hálózatokba. Harari olyan intézkedéseket javasol, mint például a cégek felelősségre vonása algoritmikus károkért (például ha a Facebook algoritmus veszélyes hazugságokat terjeszt, akkor a Facebooknak felelősségre kellene vonnia). Összefoglalva, a legfontosabb tanulság, hogy az emberiség még mindig képes irányítani az információs forradalmat, de ehhez *tudatos kollektív cselekvésre* van szükség – átláthatóság érvényesítésére, intézmények megerősítésére, és talán még a gazdasági modellek újragondolására is szükség (ahogy Harari az MI korszakában az adatadózásról és az új valutákról elmélkedik). Ez a felelősség témája a könyv ellenszere a saját riasztó helyzeteire.

Kontextus Harari tágabb munkásságában

A Nexus Harari munkásságába illeszkedik, mint a nagy történelmi történetmesélésének folytatása, valamint korábbi bestsellerei konkrét ötleteinek fejlődése:

- *Sapiensre építve: A Nexus számos fogalma* visszautal a *Sapiens*: Az emberiség rövid történetére. A *Sapiensben* Harari híresen azt állította, hogy a Homo sapiens azért hódította meg a világot, mert képesek vagyunk közös fikciókat (istenek, nemzetek, pénz stb.) alkotni és hinni benne. *A Nexus* ezt a felismerést mélyebbre mélyíti ezeknek a közös fikcióknak a *mechanikájába* – lényegében olyan információs hálózatokként kezeli őket, amelyeket időn át lehet tanulmányozni. A *Nexus korai* fejezetei (történetekről, dokumentumokról stb.) az emberi őstörténetet és történelmet a kommunikációs áttörések szemszögéből összefoglalják, amelyekről egy kritikus megjegyezte, hogy *a Nexus egyes részei "szinte teljesen megegyeznek a Sapiens-szel"* tartalomban (bár inkább információra fókuszálnak, mint biológiára vagy közgazdaságtanra). Azonban *a Nexus* nem csupán egy újrajátszás; a fókusz az információáramlásra szűkíti, ahol *a Sapiens* szélesebb körben terjedt. Mondhatnánk, hogy *a Nexus* egy tematikus folytatás: *Sapiens* elmagyarázta, hogy *a mítoszok* erőt adnak nekünk, és *Nexus* elmagyarázza, hogyan terjednek ezek a mítoszok, és mi történik, amikor a társadalom mitikus ragasztója leomlik.
- *A Homo Deus és a 21 lecke kiterjesztése*: Harari második könyve, *a Homo Deus: A Brief History of Tomorrow*, az emberiség jövőjéről spekulált, beleértve az algoritmusok és a mesterséges intelligencia felemelkedését, mint a Homo sapiens lehetséges utódjait. *A Nexus* ezeket a futurista szálakat felveszi, és a jelen valóságához alapozza őket. Például *a Homo Deus* bevezette a *dataizmus* fogalmát – azt az új ideológiát, amely mindenekelőtt a adatáramlást értékeli. A *Nexusban* Harari gyakorlatilag frissíti az adatizmust aktuális példákkal: olyan MI rendszerekkel, amelyek már az információt kurálják számunkra, a közösségi hálózatokat, amelyek adatcserét végeznek, valamint az MI által generált kultuszok vagy ideológiák lehetőségével. A *Nexus* apokaliptikus

hangulatát (az MI véget vet az emberi történelem vége, stb.) előrevetítette a *Homo Deus*, de most Harariban sokkal több anyag (például generatív mesterséges intelligencia és valós események) van rá merítve. Az Atlantic egyik kritikus megjegyezte, hogy a *Nexus* Harari "eddig legkomorabb munkája", amely "megragadó víziót kínál arról, hogyan válhat az MI katasztrófálisra", még ha nem is haladja felül a *Sapiens úttörő* minőségét. Ez tükrözi, hogyan lép tovább a *Nexus* Harari *Homo Deus* előrejelzéseinek sötét oldalára.

Harari harmadik kötete, a *21 tanulság a 21. századra*, esszékből álló gyűjtemény volt, amelyek azonnali problémákkal (terrorizmus, álhírek, technológia stb.) foglalkoztak. A *Nexus* egyetlen, a 21 Lessons központi témájának egységesebb, mélyreható vizsgálataként tekinthető: a modern világ információs válsága. A *21 lecke* című művében Harari az internet igazság aláásásáról és az MI liberális értékek kihívásáról írt. *Nexus* ezt a vitát teljes történelmi narratívává és cselekvésre való felhívássá bővíti. Különösen a *21 Lessons*-ban Harari figyelmeztetett, hogy a liberális elit gondolkodásában apokaliptikus a zavarodottság miatt válik; ironikus módon *Nexus*-ban Harari maga is az apokalipszis felé hajlik (egy kritikus ironikusan megjegyezte, hogy a *Nexus*-ban Harari "maga is liberális" riadalmas lett az emberiség végével kapcsolatos riadalmassal. Ez a viszonylag megfontolt *21 Lessons*-ról a sürgetőbb *Nexus* felé változást jelent.

- Távozások és új hangsúlyozások: Miközben a *Nexus* továbbviszi Harari jellegzetes nagykép-elemzési stílusát, új irányokat is mutat. Egy különbség a személyes érintés – Harari magában foglalja nagyapja romániai történetét, valamint esetleg más személyes vagy családi anekdotákat is, ami korábban nagyrészt hiányzik korábbi műveiből. Ez emberi elemet ad az információs hálózatok vitájához. Egy másik változás Harari kifejezetten politikai képviselője: korábbi könyvek kérdéseket vetettek fel, de nem találtak konkrét megoldásokat, míg a *Nexus* egyértelműen az AI szabályozását és a technológiai vállalatok felelősségre vonását szorgalmazza. Ez arra utal, hogy Harari megfigyelőből aktivista történészre vált, platformját felhasználva befolyásolja az MI irányításról szóló vitát.

Emellett a *Nexus* szűkebb a látókörben, mint a *Sapiens* vagy a *Homo Deus*, amelyek időben és témában is kiterjedtek. Az információs hálózatokra koncentrálva a *Nexus* feláldoz némi szélességet a mélység érdekében ezen az egyetlen témán. Néhány kritikus ezt a fókuszot kissé ismétlődőnek találta (egyikük megjegyezte, hogy *Sapiens*-nek tűnik, "az őskori emberről szóló részeket lerövidítve, és a generikus MI jereimád elhúzódot"). Másrészt a támogatók úgy látják, hogy Harari ötleteinek időszerű szintézise, amely a jelenlegi MI-felforduló pillanatra igazodott. Röviden, a *Nexus* Harari ötlettrilógiájának csúcspontja: a fikció ereje (*Sapiens*), a jövő technológiájának fenyegetése (*Homo Deus*) és a jelen infokalipszis navigálásának kihívása (*21 Lessons*) – mind egy narratívába fonva.

A szerző szándéka és kommentárja a *Nexus*hoz

Harari interjúban és esszéiben beszélt szándékairól a *Nexus*-szal, ezzel rávilágítva, miért írta most ezt a könyvet. Egy Vox interjúban elmagyarázta, hogy a könyv alapvető kérdése az: "Ha az emberek ilyen okosak, miért vagyunk mi ennyire ostobák?". Ezzel arra gondol: hogyan lehet egy faj, amely képes feltérképezni a genomot és felfedezni a galaxisokat, ugyanakkor az önpusztítás szélén állhat? Harari nyíltan elutasítja a régi mitológiai választ, miszerint "valami

eleve hibás az emberi természetben." Ehelyett tézise – és szándéka – az, hogy megmutassa, a hiba az információinkban rejlik. *"A legtöbb ember jó ember... De ha rossz információkat adsz a jó embereknek, rossz döntéseket hoznak"* – mondta a Voxnak. Így Harari célja, hogy újrírja az emberi bolondság narratíváját: nem az eredendő bűn vagy veleszületett gonoszság, hanem a *félrevezető információk* terjedése a hálózatainkon keresztül, ami félrevezet minket.

Harari azt mondta, hogy a *Nexus-t* figyelmeztetésként és útmutatóként írta a jelen pillanatra. Megjegyzi, hogy *"az emberiség történetének legmélyebb információforradalmaján"* élünk, és ahhoz, hogy megértsük a jelent (és túléljük a jövőt), meg kell értenünk az információ múltbeli forradalmait. Lényegében a *Nexus* célja, hogy történelmi perspektívát nyújtson a digitális/mesterséges intelligencia korszakára. Harari kijelentette, hogy hosszú távú nézőpontot szeretne bemutatni az olvasóknak – a kőkortól a Szilícium-völgyig –, hogy megértsék, miért olyan zavaró az MI és az internet. Az olyan jelenségek összekapcsolásával a kőkorszaki törzsi mítoszok és a Facebook-vezérelt népi társaság összekapcsolásával meg akarja mutatni, hogy az igazság és a hazugság kihívásai mindig is velünk voltak, bár a tét most sokkal magasabb.

Harari szándékának másik aspektusa, hogy nyilvános beszélgetést indítson az MI és a demokrácia metszéspontjáról. Olyan cikkekben és panelekben hangosan kifejezte, hogy a liberális demokrácia nem feltétlenül éli túl az AI hatását tudatos cselekvés nélkül. Egy interjúban Harari azt mondta, reméli, hogy a könyv arra ösztönzi az embereket (különösen a döntéshozókat), hogy *"tájékozott döntéseket hozzanak"* az MI-vel kapcsolatban, és ne adják le a felelősséget a technológiai cégekre vagy a sorsra. Ez összhangban áll aktivizmusával: Harari a *Nexus* könyvbemutatóját arra használta, hogy olyan dolgokat képviseljen, mint a globális szabályozási keretrendszer az MI számára, valamint a közösségi média algoritmusok szigorúbb felügyelete. Gyakran erkölcsi szempontból fogalmazza meg ezt – például: "ha nem tudjuk megváltoztatni a jövőt, akkor miért beszélünk róla?" – utalva arra, hogy a *Nexus* témáinak megvitatása része annak, hogy visszaszerezzük a jövőnk irányítását.

Harari számos nyilvános megjegyzést tett, amelyek összefoglalták a *Nexus* alapvető üzenetét. Egy megtörekvő mondat, amit gyakran ismétel: *"Az információ nem igazság. Az információ legfőbb dolga, hogy összeköti az embereket, és a legegyszerűbb módja a kapcsolatba a fantáziák és hazugságok."* Ezt a beszélgetésekben felerősítve Harari hangsúlyozza, hogy megcáfolja a techno-utópia elképzelését (azt az elképzelést, hogy több adat több igazságot jelent). Kifejezetten bírálta a Szilícium-völgy gondolkodásmódját, amely az információt veleszületett jónak tekinti. Célja, hogy a technológusokat és a közvéleményt arra ösztönözze, hogy szkeptikusabbak legyenek az információ minőségével kapcsolatban, ne csak a mennyiséggel.

A *Nexus* népszerűsítése közben Harari azt is megjegyezte, miért választotta a "Nexus" címet. *Nexust* összekapcsoló pontként írja le – ebben az esetben az MI válik az összes információáramlás csomópontjává. A könyv YouTube-bevezetőjében Harari feltette a kérdést: *Ha az MI válik a hatalom és a tudás új csomópontjává, vajon vajon egy mindent átfogó fikciót teremt, amelyen az emberek nem láthatnak túlról?* (Ez visszhangozza a könyvben írt mondatát arról, hogy az MI esetleg létrehozhat egy "téveszmék hálózatát", amelyet a jövő generációi talán nem tudnak áttörni.) Ezért szándéka egyértelműen óvatos: figyelmeztetni akar minket, hogy *elménket és társadalmunkat nem emberi információs ügynökökre bízunk*, és kritikusan kell gondolkodnunk erről.

Összefoglalva, Harari azért írta a *Nexus-nak*, hogy időben figyelmeztetést adjon, és hogy az olvasók történelmi keretet kapjanak a mai információválságok megértéséhez. A könyvet

egyszerre pozícionálta a jelenlegi "ostobaságunk" (félrevezetett cselekedeteink) diagnózisának, valamint *cselekvésre való felhívásként* a bölcsesség visszaszerzésére. Harari saját nyilatkozatai a felelősséget és a reményt hangsúlyozzák: az apokaliptikus forgatókönyvek ellenére ragaszkodik hozzá, hogy a történelem nem előre elhatározott, és hogy az emberek megelőzhetik a legrosszabb kimeneteket a múltból tanulva és a jövő technológiáinak szabályozásával. Ez a reményteljes megjegyzés – hogy a tudatosság változáshoz vezethet – nagyban hozzájárul ahhoz, hogy Harari ennyire foglalkozik a Nexus körüli médiával: nemcsak a világot akarja elemezni, hanem megváltoztatni azt (vagy legalábbis azt a módot, ahogyan az információról és a hatalomról gondolkodunk).

Kritikai fogadtatás

A Nexus széles körű vitát váltott ki a médiában, és vegyes kritikai fogadtatásokat kapott, a magas dicsérettől a éles kritikákig terjedt. Gyorsan kereskedelmi sikert aratott – a New York Times #1 bestsellereként debütált, és több országban is a listák élére került –, de a profi kritikusok megosztottak voltak az érdemei miatt.

Sok kritikus dicsérte a könyv ambícióját és időszerűségét. Például a Kirkus Reviews dicsérő értékelést adott a Nexusnak, azt javasolva, hogy *"az MI-ről szóló könyvek lavinájával szembenézve az olvasók jól jönnének, ha ezzel kezdenék."* Kirkus és mások értékelték Harari képességét, hogy hatalmas történelemterületeket szintetizáljon, és azokat a sürgető kortárs kérdésekhez kötözze. A Booklist (American Library Association) szintén kiadott egy "rave" kritikát, amelyben tapsolt, hogy Harari *"a történelem, filozófia, tudomány, pszichológia és politikaelmélet alapjaiból merít"*, hogy feltárja az emberi törekvések mögött rejlő információs mintákat. A Booklist kritikusa a Nexus-t *"fontosnak és időszerűnek"* nevezte... *kötelező olvasmány, hiszen túlélésünk az információ kegyén múlik.* Az ilyen pozitív kritikák kiemelik a könyv világhírűségét a bonyolult témák magyarázatában, valamint sürgető jelentőségét az MI korában.

Néhány kiadó mérsékelten pozitív vagy kiegyensúlyozott értékeléseket adott. Az Atlantic például a Nexus-t *"eddig legkomorabb munkájaként"* jellemezte, és megjegyezte, hogy *"megragadó víziót kínál arról, hogyan válhat az MI katasztrófálisra."* Daniel Immerwahr kritikus azt írta, hogy bár a könyv nem éri el a Sapiens narratív csúcst, *"jól ír a faj méretében"*, és hatékonyan közvetíti az ellenőrizetlen MI veszélyeit. Hasonlóképpen, Killian Fox egy Guardian kritikája szerint a Nexus *"magával ragadó... diagnózist és cselekvésre felhívást"*, bár *"furcsa vakfoltokkal"* (például alig említik a kapitalizmus szerepét a Big Tech-ben). Fox arra a következtetésre jutott, hogy *akár egyetértünk Harari történelmi MI-keretezésével, akár nem, nehéz nem lenyűgöznünk az aprólékos felépítése miatt.* Ez egy közismert érzést tükröz: még a szkeptikusok is csodálták Harari történetmesélését és átfogó elemzését, még ha néhány részletet vitattak is.

Másrészt a Nexus több neves kritikustól is jelentős kritikát érzett. Az Egyesült Királyságban a The Guardian tudományos újságírója, Steven Poole egy különösen éles kritikát írt (*"A napok vége?" címmel?*) Hararit azzal vádolva, hogy *"apokaliptikus pontifikáció"* van, amely *"túlnyúl a hihetlenséget"*. Poole azzal érvelt, hogy Harari szalmabáb fogalmakat állít fel (mint egy karikatúra "naiv információnézet"), csak hogy lebontja azokat, és túl sokat támaszkodik az átfogó általánosításokra. A Guardian cikke elismerte, hogy a Nexus *"sok... lenyűgöző beszélgetések"*, és hogy Harari *"kiváló narratív író"* lehet, amikor nem *"jóslatos"* üzemmódban van. Összességében azonban Poole nem találta meggyőzőnek a könyv mesterséges intelligenciáról szóló ítéleteit. Például rámutatott, hogy eddig egyetlen mesterséges

intelligencia sem hozott létre *igazán új* ötleteket vagy művészetet, ellentétben Harari javasolataival, Hararit "különösen hiszékenynek" nevezve a jelenlegi MI képességekkel kapcsolatosnak. Úgy érezte, hogy Harari nagy megoldása (az algoritmusok szabályozása, a demokrácia megerősítése) "*gyenge következtetés*" volt az apokalipszis beszédek után.

Az Egyesült Államokban a The New York Times Book Review vegyes kritikát adott. Dennis Duncan kritikus a *Nexus-t* "*hasznos, jól informált bevezetőnek*" nevezte az információs korszakhoz, de megjegyezte, hogy "*nem minden Nexus érzi magát eredetinek. Ha figyelsz a hírekre, felismered néhány történetet, amit Harari mesél.*" Más szóval, Harari példái (mint a Facebook és a QAnon) frissek és ismerősek, így a könyv egyes részei inkább ismert események összefoglalójának érződik, nem pedig friss betekintésnek. Duncan dicsérte Harari emlékezetes tisztaságát a helyzet összefoglalásában, de végül frusztrálónak találta a könyvet, valószínűleg azért, mert riasztót kelt anélkül, hogy teljesen meggyőzné az olvasót az új megoldásokról.

A Wall Street Journal negatívabb hangvételű volt. Dominic Green kritikus gúnyolta Harari formuláját, azt írva, hogy a *Nexus* olyan, mint "*egy adag történelmi anekdota... Egy csipetnyi társadalomtudomány... egy kanál találgatás, amelyet nedves receptkéreg díszít, és enyhén az apokalipszis előérzetével megporítva.*" Az eredmény, tréfásodott, "*könnyen leesik, de nem olyan tápláló, mint ahogy állítják*", így "*savanyú ízt hagy maga után.*" A könyvet a *legérdekesebbnek és leghibásabbnak* találta, amikor a jelenlegi helyzetet kezelte, ami arra utal, hogy Harari történelmi újramesélése szilárdabb volt, mint a futurisztikus jóslatai. A Washington Post is bírálta a könyvet, Justin Smith-Rui bírálta azt a "*különösen gyenge szakaszt*", ahol Harari párhuzamokat vont az állati viselkedés és az emberi mítoszok között, és azt mondta, hogy Harari nézete "*egyre zavarosabbá válik*", amikor az MI-ről beszél. A Post szerint a *Nexus*, Harari múltbeli sikereire támaszkodva, úgy érződött, mint egy "nagy költségvetésű produkció, több kézzel készülő – utalva arra, hogy talán túlgyártott vagy kissé személytelen.

Az akadémiai és szakértői kommentárok a *Nexusról* visszhangozták ezeket a vegyes reakciókat. Sok tudós tisztelt Harari történetmesélését, de kételkedik bizonyos területeken szerzett szakértelmében. Például néhány MI-kutató és számítástechnikus feldühült Harari drámai kijelentéseire az MI-ről. Az online fórumokon kritikusok megjegyezték, hogy "*hamis állításokat tesz az MI-ről, miközben egyáltalán nem is számítástechnikus.*" Azt állítják, hogy Harari néha a legrosszabb forgatókönyveket festi meg anélkül, hogy értené a technikai korlátokat – lényegében azzal vádolják, hogy riasztó kívülálló az MI-vitában. Hasonlóképpen, a történészek és társadalomtudósok vitatják Harari túlzott általánosítási hajlamát. Harari régóta célpontja a szakosodottabb tudósoknak, akik úgy érzik, árnyaltságot áldoz fel egy nagy narratíváért. Ahogy egy hozzászóló (antropológus) nyersen megfogalmazta: "*A Sapiens tudományosan rendkívül hibás, és sok antropológus széttépte. Ez a pop tudományi hülyeség.*" Hasonló kritika érvényes a *Nexusra* is: hogy lekötő, de nem szigorú. Például a Reason Magazine kritikája Hararit hibáztatta azért, mert összetett történelmi eseményeket "*friss epizódokba*" zsúfol, és "*nem illő részleteket hagy el*", amelyek nem illeszkednek a téziséhez. A kritikus, Corbin Barthold konkrét példákat hozott arra, hogy Harari története túlegyszerűsített vagy vitathatóan félrevezető volt – például Harari állítása, miszerint a Római Köztársaság a nem megfelelő információs hálózatok miatt bukott meg, amelyet az Reason szerint figyelmen kívül hagyott gazdasági és politikai tényezők (és még azt is, hogy Julius Caesar *hatékonyan használta* az információs hálózatokat a publikált kommentárjai révén). Ez a fajta kritika azt sugallja, hogy az akadémikusok Harari történelmi érveit érdekesnek találják, de nem teljesen meggyőzőek, ha alaposan megvizsgálják őket.

Összefoglalva, a Nexus kritikai fogadtatása polarizált, de hatásos jellemzi. A rajongók dicsérik Harari lebilincselő prózáját, a nagy kép szintézisét, valamint a könyv sürgető jelentőségét az MI és a társadalom vitájának kiindulásában. A kritikusok azonban bírálják Harari *szenzációhajhászságát és technikai mélységhiányát*, valamint a széles kijelentések hajlamát is, amelyek néha túlegyszerűsítik a valóságot. A kritikák ellenére a *Nexus* sikeresen ösztönözte a széles körű vitát – ami, mondhatnánk, hogy Harari célja eléri, hogy elgondolkodtatja az embereket ezekről a kérdésekről. Ahogy a *New York Times* kritikusa elismerte, a könyv legjobb esetben *"emlékezetes tisztasággal foglalja össze a jelenlegi helyzetet"*, kiindulópontot nyújtva az olvasóknak (és a kritikusoknak) az információs dilemmák feldolgozásához a 21. században.

Kiadás és globális elérhetőség

A *Nexus* először 2024. szeptember 10-én jelent meg az Egyesült Királyságban a Fern Press kiadásában, amely a Penguin Random House egyik kiadója. Az Egyesült Államokban és Kanadában nagyjából ugyanebben az időben adta ki a Penguin Random House (Spiegel & Grau vagy Random House kiadó alatt), így gyakorlatilag egyidejű nemzetközi megjelenés is volt. A könyv angol nyelvű keménykötésű kiadásában 528 oldalas kiadás. A Harari világszintű követőtábor *miatt a Nexus* gyorsan globális terjesztést ért el. Harari hivatalos oldala szerint a könyvet rövid idő alatt közel 50 nyelvre fordították le – ami bizonyítja a nemzetközi érdeklődést művei iránt. Főbb fordítások közé tartoznak a spanyol, francia, német, kínai, japán, héber, arab és még sok más, amelyeket gyakran 2024 végére vagy 2025-re adnak ki, hogy kihasználják a könyv lendületét.

Megjelenésekor a *Nexus* azonnal bestsellerré vált számos országban. Az Egyesült Államokban (a NYT bestseller listáján a #1-es helyet érte), az Egyesült Királyságban (ahol a Sunday Times bestseller első helyén szerepelt), valamint más piacokon (például Kanadában, Németországban, Braziliában, Olaszországban, Spanyolországban és Japánban) a nem fikciós listák élére került. Harari csapata büszkén jelentette meg az Instagram-ban, hogy a NYT listáján a #2. helyen debütált (valószínűleg az első héten), és a #3. helyen az Egyesült Királyság Sunday Times listáján. Ez a kereskedelmi siker tükrözte Harari korábbi könyveit, és biztosította, hogy a *Nexus* széles körben elérhető legyen a könyvesboltokban világszerte.

A könyv több formátumban is elérhető: hagyományos nyomtatott, e-könyv és hangoskönyv. Az audiokönyv (amelyet angolul beszélt Derek Perkins, ha a korábbi gyakorlatot követi) egyszerre jelent meg, és az e-könyv kiadások minden nagyobb platformon elérhetők. Harari weboldala linkeket kínál a *Nexus* vásárlásához különböző régiókban és nyelvekben, ami egy összehangolt globális kiadást tükröz. 2025-re a *Nexus* a legtöbb nagy könyvtárban megtalálható, és gyakori látvány volt a repülőtéren könyvesboltokban, ami széles körű elérhetőségét tükrözi.

A kiadók tekintetében: az Egyesült Államokban a *Nexus-t* a Spiegel & Grau (ugyanaz a Sapiens kiadó az Egyesült Államokban) adta ki a Penguin Random House alatt, az Egyesült Királyságban pedig Jonathan Cape (Penguin) vagy az új Fern Press kiadó. Néhány régióknak voltak speciális kiadói (például Indiában a könyvet a Penguin helyi ápolón keresztül terjesztették). A koordinált időzítés segített abban, hogy Harari 2024 őszi nagy új megjelenéseként reklámozza.

A gyors fordítás ~50 nyelvre azt mutatja, hogy Harari üzenete várhatóan különböző kultúrák között is rezonálni fog. Valóban, az információ és félretájékoztatás témái világszinten is relevánsak, és Ázsiából Európáig a kiadók gyorsan megszerezték a jogokat. 2025-re a *Nexus* nemcsak nyugati nyelveken volt elérhető, hanem olyan nyelveken is, mint kínai, japán, koreai,

arab, héber, török, orosz, lengyel és több indiai nyelven, többek között. Harari csapata a Sapienshipnél (az ő szervezeténél) együttműködött nemzetközi kiadókkal, hogy viszonylag gyors fordítást biztosítsanak a fordított kiadások elterjedésében.

Egy érdekes kiadvány részlete, hogy Harari átfogó végszövegeket és hivatkozásokat választott a *Nexus* számára. A hivatalos oldal még letölthető PDF-et is tartalmaz a könyv hivatkozásairól. Ez arra utal, hogy a források dokumentálására törekednek a könyv potenciálisan vitatott állításai miatt, és talán válaszokat adnak azoknak a korábbi kritikusoknak, akik megkérdőjelezték Harari ténybeli pontosságát. A legtöbb nyomtatott kiadás tartalmazza ezeket a végjegyzeteket (amelyek több tucat oldalt ölelnek, és a források széles körét mutatja, az akadémiai cikkektől a hírcikkekig 2023–2024-ig).

Összefoglalva, a *Nexus* 2024 szeptemberében nagy globális felhajtással jelent meg a Penguin Random House néven, és azóta világszerte elérhetővé vált számos nyelven. Kiadása jelentős esemény volt a kiadói naptárban, amelyet egy globális marketingkampány és széles körű terjesztés támogatott – biztosítva, hogy szinte bárhol hozzáférjenek Harari legújabb információihoz az információs hálózatok jövőjéről.

Jelentős viták és közbeszéd

Provokatív tartalma miatt a *Nexus* nyilvános vitát és néhány vitát indított a hivatalos könyvkritikákon túl. Az egyik fő vitakérdés Harari MI-vel kapcsolatos álláspontja és hitelessége volt ezen a területen. Néhány technológiai szakértő és kommentátor visszavonta Harari romboló előrejelzéseit, azzal érvelve, hogy túllép azzal, hogy túllép a szakértelmén kívüli spekulációkkal. Például a Redditen és a közösségi médiában vita zajlott Harari MI-vel kapcsolatos megértéséről, olyanok voltak, mint "*Hararinak nincs munkaismerete a kódolásról vagy az AI megvalósításról; azok, akik a leghamisabb állításokat tesznek az MI-ről, azok azok, akik nem ismerik a technikai részleteket.*". Ez egy szélesebb vitához kapcsolódik: a Szilícium-völgy optimizmusa és az intellektuális pesszimizmus ütközésére. A technológia támogatói (köztük olyan személyek, mint Mark Andreessen, akinek "Miért fog MI megmenteni a világot" című esszéjét a *Nexus implicit* módon megkérdőjelezi) azt állítják, hogy Harari félelmet kelt az MI-vel kapcsolatban. Ezzel szemben Harari és mások arra figyelmeztetnek, hogy az ilyen optimizmus figyelmen kívül hagyja a valódi veszélyeket. Ez a vita véleménycikkeken és online fórumokon zajlott, így a *Nexus* gyakorlatilag referenciapont lett arról, hogy az MI egzisztenciális fenyegetést jelent-e vagy csak egy újabb technológiai eszköz.

Egy másik vita alakult ki Harari történelmi állításai körül. Történészek és kommentátorok megkérdőjelezték Harari néhány példáját és következtetését, ami élénk vitákat eredményezett blogokban és magazinokban. Például Harari állítása, miszerint Gutenberg nyomdáját némileg felelősségre vonják a boszorkányüldözésért (boszorkánykézikönyvek terjesztésével), némi gúnyos fogadtatásban részesült – a kritikusok szerint ez figyelmen kívül hagyja a nyomda pozitív szerepét a tudományban, és hogy Harari ok-okozati összefüggése túl leegyszerűsített. Hasonlóképpen, a Római Köztársaság bukásának információs szemüvegen keresztüli értelmezése (amely szerint Róma nem rendelkezik a tömegdemokráciához szükséges információs infrastruktúrával) vitatott volt, ahogy azt az Reason cáfolata is mutatja, amelyben olyan ellenpéldákat emelt ki, mint Caesar hírlevelek használata. Ezek inkább tudományos viták, de közéleti diskurzusba is áttértek, miközben a történelem fórumokon és a Twitteren Harari érveit bontották szét. Lényegében Harari hajlama a merész történelmi analógiákra (az egyik védjegye) beszédtemává vált – egyesek üdvözölték a friss nézőpontot, mások azzal vádolták, hogy "a tényeket a narratívájához igazítja".

A Nexus *apokaliptikus hangulata* is felkeltette a nyilvános vitát. A könyv rajongói Harari figyelmeztetéseit az MI-vel kapcsolatos meggyőzőnek és fontosnak találták, gyakran idézve a saját helyzeteit, amikor a valós AI híreiről beszélgettek (például a GPT modellek vagy deepfake-ek fejlesztéseiről). Másrészt a kritikusok Hararit ítélőnek nevezték. Figyelemre méltó, hogy a Nexus körülbelül akkor jelent meg, amikor több technológiai híresség is kifejezte az MI félelmeket (2023–2024-ben nyílt levelek írtak MI szünetekre stb.), így Harari nézeteit néha egy kicsit a szélesebb "MI pánik" narratívába sorolták. Ez gondolatmenetekhez vezetett, amelyek azt kérdezték, hogy túlzó-e az MI riadóssága, és Harari-t gyakran említik. Például egy vélemény szerző a *The Times*-ban (brit) kigúnyolta Harari "ragaszkodóságát, hogy fajunk a század végére kihál", mondván, hogy az olvasók megbocsáthatók, ha nem veszik túl komolyan ezt a merész jóslatok alapján. Ez egy kis vitává vált: vajon Harari valóban azt mondta, hogy az emberiség 2100-ra kihál? A könyvben egy legrosszabb esetet sugall, amikor a *homo sapiens*, ahogyan ismerjük, talán nem éli túl (akár mesterséges intelligencia átalakulása vagy katasztrófa révén), de egyesek ezt túlzásnak érezték hatás szempontjából. Harari támogatói azt állítják, hogy ebben nem "biztos", csak figyelmeztet, mi *történhet*, ha meggondolatlanok vagyunk.

Vita folyt továbbá Harari közéleti értelmiségi szerepéről is. Néhány kommentátor azzal vádolja, hogy "guru", aki túlzásba viszi a tömeg kedvéért. Az online kritika egy szála a Nexus-t (és Harari munkáját általában) "ál-intellektuális szemétnak, amit az ökoszisztémájában éleslátónak tartanak". Másrészt sok olvasó – köztük jelentős személyek is – nyilvánosan dicsérték Harari bátorságát, hogy nagy kérdéseket kezeljen. Például Tom Hanks színész lelkesen ajánlotta a *Nexust*, mondván: "Ha idén csak egy ismeretterjesztő könyvet olvasol, fontold meg ezt, hogy jobban megismerd történelmünket, és tisztában legyél a következő döntéseinknek.". Hasonlóképpen, Stephen Fry a könyvet "hatalmas, gondolkodtatónak" nevezte... izgalmas és hátborzongató," még a vezetőket is arra biztatva, hogy olvassák el. Az ilyen kedvelt közszereplők támogatásai pozitív visszhangot keltettek, és ellensúlyozták az akadémiai ellenállást. A diskurzus egy kis ironiát is megjegyez: Harari, aki figyelmeztet a híresség-szerű "prófétákra" és a tömeges téveszmékre, maga is kvázi-guru státuszt szerzett, hatalmas követőtáborral. Ez nem került el a kommentátorok figyelmén, akik közül néhányan gúnyosan "szentírásnak" nevezik a technológiától féлдő osztály számára. Harari pedig elismerte, hogy *finom vonalon jár* – egyszerűsíteni akarja a bonyolult ötleteket, hogy elérje a közönséget, még a kritika kockázata ellenére is. Ahogy egy Reddit-felhasználó (aki Hararit védte) fogalmazott: "A Nexus nem akadémiai műnek számít... Ez a hétköznapi popfilozófia, hogy az emberek másképp gondolkodjanak az információról és a mi kapcsolatunkról vele.". Ez az érzés megragadja a közbeszéd lényegét: vajon Harari közszolgálatot tesz azzal, hogy népszerűsíti ezeket a kérdéseket, vagy inkább hígítja azokat?

A jelentős dicséretet tekintve, a bestseller státuszon és a hírességek elismerésén túl a Nexus megérintette a közönséges olvasókat, akik megosztották tapasztalataikat online. Sok olvasó olyan platformokon, mint a Goodreads és a Reddit, azt mondta, hogy a könyv tudatosabbá tette őket arra, hogyan fogyasztanak információt. Néhányan a "naiv információ nézet" fogalmát idézték (amit Harari kritizál), és elismerték, hogy nem is értették, hogy *több információ nem automatikusan jobb*. A Reddit r/YuvalNoahHarari oldalán egy felhasználó összefoglalta a fő tanulságot: az emberiség önkorrektív mechanizmusai (mint például a tudomány és a demokrácia) kulcsfontosságúak, és a Nexus megerősítette, miért kell megvédeni és fejleszteni azokat, akik az MI korában vannak. Azonban még a rajongói közösségekben is vita volt – egy olvasó panaszkodott: "Már nem számoltam számom, hányszor mondta azt a naiv nézetet" – úgy éreztem, mintha gyerekekkel bánnának", ami arra utal, hogy Harari pedagógiai hangvétele nem illik mindenkihez.

Nem alakultak ki feltűnő *botrányok* vagy "törlések" a *Nexus* körül – a viták szellemi jellegűek voltak. A nyilvános diskurzus egyik figyelemre méltó pillanata Harari részvétele volt a 2023/24-es nyílt levélkampányokban az AI kockázatokról. Harari társszerzője az MI-re figyelmeztető leveleket (köztük egyet, amely az AI hatását a nukleáris háborús kockázattal hasonlította össze), és véleménycikkeket írt olyan kiadványokban, mint *a The Economist*, a *Nexus* megjelenése körül. Ez vitákhoz vezetett hírműsorokban, ahol Harari pontjait vitatták – néha az AI fejlesztők ellenállásával. Egy televíziós panelen például egy MI cég vezérigazgatója azt állította, hogy Harari víziója túl disztópikus, amire egy másik szakértő azt válaszolta, hogy a Harari értékes szolgáltatást nyújt azzal, hogy a legrosszabb eseteket képzele el, hogy elkerüljük őket. Így *a Nexus* Hararit kulcsfontosságú hangként pozícionálta az MI globális párbeszédében – ez a szerep természetesen dicsérettel és vitákkal jár.

Összefoglalva, a *Nexus* körüli nyilvános diskurzus élénk volt. A könyv sikeresen teljesítette azt a célt, hogy az embereket az információ, technológia és a hatalom metszéspontjáról beszéljenek. A rajongók dicsérik Hararit, amiért riasztó és ösztönöző elmélkedést (gyakran idézve fülbemászó mondatait és helyzeteit), míg a kritikusok megkérdőjelezzik következtetéseit és tekintélyét (néha éles nyelvű megjegyzésekkel a stílusáról és pontosságáról). Ez a dinamika tükrözi a könyv témáját: az igazság és bölcsesség megfelismerésének küzdelmét az információk és vélemények kakofóniájában. Ebben az értelemben *a Nexus* egyfajta Rorschach-teszté vált – az emberek reakciói gyakran annyit mondanak a saját információs filozófiájukról, mint magáról a könyvről.

Stílus, hangulat és hozzáférhetőség

Harari írásstílus *a Nexusban* ismerős lesz a korábbi művei olvasóinak: lebilincselő, átfogó, és a nagyközönség felé szól. Érzéke van abban, hogy összetett gondolatokat világos, könnyen érthető prózában mutassa be – gyakran anekdotákon és analógiákon keresztül. A kritikusok általában egyetértenek abban, hogy Harari *nagyon olvasható narratív stílusban ír*, még akkor is, ha évezredek vagy technikai témákat fed le. A *Vox* írója, Sean Illing megjegyezte, hogy mint minden Harari könyve, *a Nexus "rengeteg területet fed le, de ezt könnyen megemészthetően csinálja."* Valóban, Harari történetmesélési technikákat – drámai történelmi jeleneteket, gondolatkísérleteket és retorikai kérdéseket – használ, hogy az olvasót lekötve tartsa. A fejezetek logikus sorrendben folynak (főként időrendi sorrendben), és a fejezeteken belül gyakran egy feltűnő történelmi epizóddal nyit (boszorkányper, sztálinista tisztogatás stb.), hogy szemléltesse a témát, majd kifelé terjeszkedik.

A Nexus hangulata azonban jellegzetesen sürgetőbb és komolyabb, mint Harari korábbi könyvei. Ahogy több kritikus is megjegyezte, ez Harari eddigi legapokaliptikusabb vagy legóvatosabb munkája. Míg a *Sapiens* gyakran szellemes, szinte játékos hangvételő volt az emberi hibák leírásakor, *a Nexus* komolyabb. Harari aggodalma információs ökoszisztémánk esetleges összeomlása miatt néha komor hangvételt kölcsönöz – olyan módon beszél az *"emberi történelem végéről"* és a *"civilizációs öngyilkosságról"*, amelyek felkelti a figyelmet. A hangnemet figyelmeztető vagy prófétikusnak is lehetne leírni: Harari riadót ad, és nem riadja vissza a drámai nyelvezetet a megtegye. Ennek vannak előnyei és hátrányai: érzelmi súlyt ad a könyvnek, de néhány olvasó túlterheltnak vagy túlzónak találta (ezek az olvasók például "pontifikáció" vagy "jeremiad" szavakat használhatnak, ahogy azt a negatív kritikákban láthatják).

A sötétebb hangvétel ellenére Harari írása továbbra is nagyon hozzáférhető. Nem feltételezi az olvasótól a történelmi vagy technológiai előzetes ismereteket. A kulcsfogalmak (mint például

a Turing-teszt, vagy hogy mi az ideghálózat) egyszerű nyelven vagy metaforákon keresztül magyarázzák el. Történelmi epizódok bemutatásakor gyakran éppen annyi kontextust ad, hogy egy laikus megértse a jelentőségét. A könyv rövid alfejezetekre van felépítve fejezeteken belül, ami segít az olvashatóságban – darabonként fel lehet emészteni. Harari mondatai általában világosak és ütősek; bőven használja retorikai kérdéseket, hogy kvázi-beszélgetésbe vonja az olvasót. Például megkérdezheti: "Mi történik, ha senki sem tudja már megmondani, mi az igaz?", majd válaszolhat. Ez a technika, valamint a kulcsszavak ismétlése, megerősíti a főbb pontokat.

Az ismétlődésről beszélve: az olvasók egyik stílusbeli furcsasága Harari gyakori "*a naiv nézet*" kifejezés használata (az információval kapcsolatban). Ismételten is felhossa ezt az érvei ellenpontjaként. Néhány olvasó pedagógiaileg hasznosnak találta ezt – hangsúlyozva a naiv optimizmus és az informált realizmus közötti különbséget. Mások lekezelőnek vagy monotonnak találták ezt ("olyan, mintha gyerekként bánnak vele" – mondta egy hozzászóló). Ez rámutat, hogy Harari hangvétele néha didaktikus irányba hajlik. Hiszen végzettségében előadó, és *a Nexusban* időnként professzori módba csúszik, megmondja az olvasónak, mit gondoljon egy adott példáról, mielőtt továbblép a következőre. A legtöbb általános olvasó azonban inkább tisztázónak találja ezt a stílust, mint lekezelőnek, mert lépésről lépésre bontja le az érveket.

Harari példák és analógiák használata a könyv hozzáférhetőségét is meghatározza. Összehasonlíthatna egy mesterséges intelligencia algoritmust egy vallási prófétával, vagy az internet adatáradatát a "tűzoltó csőből iváshoz" hasonlíthatja (hogy egy gyakori metaforával használja). Ezek az eszközök kézzelfoghatóvá teszik az elvont kérdéseket. A *Nexus című művében* élénk képeket használ – például egy MI-t "idegen intelligenciaként" ír le, vagy "Szilíciumfüggöny" részfalakat képzel el –, amelyek az olvasó emlékezetében megmaradnak. Kritikusok, mint Killian Fox, megjegyzék, hogy akár egyetértünk egyet, akár nem, Harari kerete "*aprólékos*", és *világosan "építi"* fel az érvest. A Guardian pozitív kritikája azt is megemlítette, hogy a könyv "*elsősorban diagnózisként és cselekvésre felhívásként működik*", és ezekkel a kifejezésekkel "széles körben sikeres", utalva arra, hogy meggyőző írásként eltalálja a helyét.

Közönség szempontjából *a Nexus* egyértelműen a nagyközönségnek van írva, nem a szakértőknek. Szinte semmilyen zsargont nem tartalmaz magyarázat nélkül. Még gépi tanulás vagy történelmi események esetén is Harari általában egy mondatot ad a magyarázathoz. Például AlphaGo győzelmét egy Go bajnok felett elmeséli anélkül, hogy feltételezi, hogy az olvasó tud róla, hanem úgy mesél el, mint egy történetet. Ennek a széles körű megközelítésnek az árnyaja, hogy a témákban jártas olvasók a tartalom egyes részeit alapvetőnek vagy már ismertnek találhatják (ahogy a NYT bírálója is megjegyezte – sok példa felismerhető a hírkövetők számára). De Harari célja egyértelműen az, hogy széles közönséget tájékoztasson és riasztson, nem pedig új kutatásokat mutasson be az akadémikusok előtt.

Stílusban a könyv a népszerű tudomány és a történelem szintézise szellemében halad. Felépítésében összehasonlítható olyan szerzők műveivel, mint Jared Diamond vagy Steven Pinker (nagy ötlet, amit korszakok különböző esettanulmányai is alátámasztanak). Azonban Harari hangja inkább narratív jellegű, és néha talán filozofikusabb. Nem riad vissza a nagy kijelentésektől és erkölcsi ítéletektől, ami magabiztos (egyesek szerint túlmagabiztos) hangvételt kölcsönöz a prózának. Egy Wall Street Journal részlet viccelődött, hogy Harari "*az apokalipszis előérzetei*" díszítéssel szolgálja fel az ötleteit – lényegében azt mondja, hogy képes drámai lenni. Mégis, még az a kritika is valószínűleg elismeri, hogy a könyv folyamatosan

lapozza az olvasót, mert Harari tudja, hogyan kell problémát felvetni, majd egy történelmi anekdotát ugratni, ami megvilágítja azt.

Harari hangnemének egyik figyelemre méltó eleme a *Nexusban* a távolságtartás és a szenvedély kölcsönhatása. A történelmi narratíva nagy részében hűvös, szinte klinikai távolságtartással ír – például tényyszerűen leírva, hogyan tette lehetővé a bürokrácia a népirtást. De amikor a jelenbe/jövőbe vált, a szenvedélye megmutatkozik. Mintha a történész hátrálna, és az aggódó polgár lépne előre. Ez stílusok keverékéhez vezet: részben informatív krónika, részben szenvedélyes véleménycikk. Néhány kritikus nem tetszett ez a keverék, mert úgy érezték, hogy a véleménycikk túl egyoldalú vagy spekulatív volt. Mások értékelték, hogy a könyvnek normatív álláspontja van (azaz valamit érvel, nem csak leír). Ez kissé eltér a *Nexus-tól*, mint a *Sapiens*, amely nagyrészt elkerülte a kifejezett recepteket. Itt Harari meg akar győzni, és ez a sürgősség átszivárog a stílusba.

Összefoglalva, a *Nexus* megőrzi Harari jellegzetes, hozzáférhető, történetvezérelt stílusát, biztosítva, hogy a nem szakértők is olvashatóak számára is olvasható legyen, és bevonja a narratív folyamatot. A hangvétel sürgetőbb és óvatosabb, mint korábbi művei – néha riadalmassá vált, ami megosztotta az olvasókat. A stílust a kritikusok "lebilincselő és feltűnően eredeti" és kissé "prédikálónak" nevezték, de tagadhatatlanul sikerül olyan módon közvetíteni az összetett gondolatokat az információról, technológiáról és az emberi viselkedésről, hogy ezrek olvasói megértsék őket. Harari megközelítését a *Nexusban* egy kritika foglalhatja össze: "*Harari narratívája lebilincselő, és a keretezése lenyűgözően eredeti*" – ezt a lebilincselő narratívát használja arra, hogy sürgős üzenetet közvetítsen, és a *Nexus-t* legalább egyszerre manifesztóvá teszi, mint történelemmé, olyan nyelven írva, amit szinte bárki megért.

Kapcsolódó média és események

A *Nexus megjelenése* környékén és azóta Yuval Noah Harari aktív szerepet töltött be különféle médiumokban, előadásokban és beszélgetésekben, hogy népszerűsítse a könyv ötleteit és kapcsolatot teremtsen a közönséggel világszerte. Ez a többplatformos jelenlét társként szolgált a könyvhöz, lehetővé téve Harari számára, hogy tovább magyarázza és vitatta a *Nexus* fogalmait.

Már a megjelenésekor Harari interjúkat készített nagy médiumokkal. Egy nagy figyelmet kapott példa erre a Vox "The Gray Area" podcastjában való szereplése, ahol a *Nexusról* beszélgetett Sean Illing műsorvezetővel (a beszélgetés címe: "*Yuval Noah Harari, hogy a demokrácia és az MI együtt létezhet-e*"). Ez az interjú, amelyet 2024 szeptemberének végén jelent meg, mélyen foglalkozik a könyv fő érveivel az információtúlterheléssel és az MI demokráciára gyakorolt fenyegetésével kapcsolatban. Harari idézetei ebből a podcastból – például a viccének: "Ha az emberek ennyire bölcsek, miért vagyunk technológiai öngyilkosság szélén?" – cikkekben és közösségi médiában is megosztották, hatékonyan erősítve a *Nexus* témáit. A Vox-interjú egyszerre volt promóciós szereplés és intellektuális vita, és Harari olyan kérdésekre válaszolt, amelyek a szélsőségesebb forgatókönyveivel kapcsolatos kihívásokkal foglalkoztatták, így a hallgatók a könyvön túl is több betekintést adtak. (A podcast formátum lehetővé tette Hararinak, hogy tisztázza a finomságokat, például elismerje a jövővel kapcsolatos bizonytalanságot, ami néha elveszik az oldalon.)

Harari több tévé- és videóplatformon is megjelent. A *Nexus* megjelenése utáni hónapokban interjúkban szerepelt olyan csatornákon, mint az MSNBC, a CNN és a BBC. Például 2025 júniusában az MSNBC "*Morning Joe*" című műsorában szerepelt, ahol a kortárs politikát

("Trump 'középkori' politikája") a *Nexus* szemszögéből értelmezte, és figyelmeztetett az MI által vezérelt félretájékoztatási csalásokra. Ez közvetlenül a *Nexus* elemzését a jelenlegi eseményekről szóló beszélgetések (például a választási félretájékoztatás) bevezette. Harari szintén hosszú interjút adott a Wirednek ("*A nagy interjú: Yuval Noah Harari az emberiség, mesterséges intelligencia és információ jövőjéről*" címmel) 2025 májusában), ahol válaszolt arra a kérdésre, hogy az MI miként változtathatja meg a kreativitást és kultúrát – ezzel hatékonyan bővítve a *Nexus* II. fejezetét. Az ilyen médiamegjelenések gyakran "bónusz tartalomként" szolgáltak azok számára, akik érdeklődtek a *Nexus* iránt, mivel Harari néha olyan példákat vagy kifejezéseket hozott fel, amelyek nem találhatók meg a könyvben.

Ezen felül Harari egyfajta előadás- és előadókörútra indult, amely a *Nexushoz* kapcsolódik. 2024 októberében, a megjelenés után nem sokkal a londoni How To Akadémián tartott előadást, amelyben a könyv kulcsponjtait bontotta le, és a közönség számára is kérdéseket adott fel. Hasonlóképpen, előadásokat tartott irodalmi fesztiválokon és rendezvényeken (például egy virtuális eseményen a Commonwealth Clubban, valamint élő beszélgetésen a Royal Institutionban). Ezek az előadások gyakran olyan címeket viseltek, mint az "Az információ és az MI jövője", és lényegében Harari-k voltak a moderátorokkal folytatott beszélgetések a *Nexusról*. Interaktív lehetőséget biztosítottak az olvasók számára – a közönség tagjai kérdéseket tettek fel arról, mit tudnak tenni az egyének, vagy mennyire optimista Harari, stb., így tisztázhatta, hogy még mindig van reménye, ha lépéseket tesznek (ami néha elveszik a könyv sötétségében).

Harari saját YouTube-csatornája és weboldala néhány anyagot is tartalmazott. Egy figyelemre méltó feltöltés volt a "*Yuval Noah Harari bemutatja a 'Nexus'-t*", egy rövid videó, amelyben Harari könnyen érthető módon összefoglalja a könyv alapvetését – hogy példátlan pillanatban vagyunk, ahol tanulnunk kell a történelem információforradalmaiból, hogy eligazodjunk az MI forradalom során. Fülbemászó mondatokat használ (szövegekben és interjúkban), például az MI összehasonlítása az "idegen intelligenciával", és hangsúlyozza, hogy a mítoszok egyszerre voltak a szupererőnk és az Achilles-sarkunk. Ez a videó, amelyet a közösségi médiában osztottak meg, gyors összefoglaló volt azok számára, akik még nem olvasták a könyvet, és segített érdeklődést kelteni.

Továbbá Harari néhány speciális párbeszédet is folytatott, amelyek a *Nexus* tartalmához kapcsolódnak. Például nyilvános beszélgetést folytatott Mustafa Suleymannal (a DeepMind társalapítójával és a *The Coming Wave* szerzőjével), mivel Suleyman a Nexust dicséri. Megvitatták az MI politika és a történelem metszéspontját, hatékonyan hidalva Harari ötleteit egy technológus nézőponttal. Harari podcastokban és YouTube műsorokban is szerepelt, nem pedig a tipikus könyvturnépályákon: 2025 végén a Big Think YouTube csatornán is fellépett, hogy beszéljen arról, hogy "*Miért dőlnek a fejlett társadalmak tömeges téveszmedésnek*", amely közvetlenül a *Nexus* történelmi példáiból merít. Részt vett olyan beszélgetésekben, mint a "*News Agents*" podcast, sőt egy buddhista kolostor webcastja is ("*Homo Deus találkozik a Dhamma: Információs diéta, mesterséges intelligencia és az ébredés nexusa*" 2025 májusában). Mindegyikben Harari a beszélgetést a közönséghez igazította – például elmagyarázta a szerzetesi közönségnek, hogyan védheti a mentális fegyelem (egyfajta információs diéta) a digitális félretájékoztatástól, ami visszakapcsolódik a *Nexus* témájához, amely az egészségtelen információs hálózatok elleni ellenállásra utal.

Online a *Nexus* sok olvasói beszélgetést és közösségi részvételt is indított el. A Redditen a dedikált r/YuvalNoahHarari oldal tagjai "kulcsfontosságú tapasztalatok" összefoglalóit posztoltak a *Nexus* oldaláról, és másokat is megkérdeztek Harari állításairól. Ezek a szálak néha

úgy olvashatók, mint a könyv kiterjesztései – például a felhasználók vitatkoznak Harari nézetéről, miszerint az információ többsége szemét, és csak kapcsolódáshoz hasznos, nem pedig igazsághoz. Az, hogy az olvasók a fórumokon fejezeteket gyűjtenek felsorolásokra, azt mutatja, hogyan terjedtek a könyv ötletei és saját életet öltöttek. Harari csapata időnként a közösségi médiában is interakcióba lépett, tisztázva a pontokat vagy megosztva véleményeket.

Harari saját platformja, a Sapienship, a *Nexus-t* használta ugródeszkeként a szélesebb körű kampányokhoz. A Sapienship legalább egy panelt szervezett az MI és a demokrácia témájában (amelyben Harari szakértőkkel együtt vett részt), hogy egybeessen a könyv témáival. A cél az volt, hogy a könyv üzenetét konkrét politikai vitává alakítsák – például az algoritmikus átláthatóság gyakorlati megvalósításának vizsgálata.

Összefoglalva, a *Nexus*t számos kapcsolódó média és esemény támogatta és kiegészítette:

- Interjúk és podcastok: Vox *Gray Area*, Wired *Big Interview*, The Economist podcast, Sam Harris podcastja (ahol Harari az MI-ről beszélt), többek között.
- Televízió: Hírműsorokban (CNN, MSNBC, BBC) szerepelnek, ahol az MI szabályozásáról beszélnek, gyakran hivatkozva a *Nexus* ötleteire a jelenlegi kontextusban.
- Élő beszélgetések: How To Academy előadás, irodalmi fesztivál kérdések és válaszok, akadémiái intézmények előadásai (Harari olyan helyeken beszélt, mint a Pekingi Egyetem digitális korszak kihívásairól).
- YouTube és online videók: Harari bevezető a *Nexus*-hoz, Big Think videók, együttműködések oktató YouTube-csatornákkal (például egy Kurzgesagt videó a "mítosz veszélyéről", ahol Harari hozzájárulása látható – elméletileg).
- Közösségi média és közösség: Reddit AMA-k (Harari 2025-ben készített egy AMA-t a *Nexus* témákra fókuszálva), Twitter-szálak, amelyek összefoglalták az álláspontjait (néhányik víruszerű lett, például egy "10 módja annak, hogy az MI alááshatja a demokráciát" című téma, amit a könyvből lepártoltak).
- Kísérő cikkek: Harari a megjelenés körül írt vagy közreműködött véleménycikkekben a The Economistben és a Time Magazinban, lényegében összefoglalva a kulcsfontosságú érveket. Az Economist egyik írása, amelynek címe "*Az információs háborúk hamarosan rosszabbodnak*" közvetlenül a *Nexus*ból merített, tömörített narratívát adva a könyv állításairól az adott kiadvány olvasóinak.

Mind ezek a kapcsolódó médiumok segítettek párbeszédet teremteni a *Nexus* körül. Az olvasók nemcsak a könyvet fogyaszthatták, hanem valós időben is láthatták, ahogy Harari védi és beszél róla. Ez a többszörös jelenlét megerősíti Harari álláspontjait, és egyben ellenpontokat is bemutat, amelyeket néha elismer ezekben a fórumokban. Bizonyos értelemben ez a *Nexus* önkorrektív felhívásának kiterjesztése: Harari kiterjesztette az ötleteit, és a média által kidolgozott ellenállás felé vette részt, megtestesítve azt a fajta diskurzust, amit képvisel.

Végül érdemes megjegyezni, hogy Harari könyvkörútja globális volt. Eseményeket vagy médiaszekciókat tartott az Egyesült Államokban, az Egyesült Királyságban, Európában, Indiában és gyakorlatilag más régiókban. Bárhová is ment, a helyi média gyakran kötötte a *Nexus* ötleteit a belpolitikai kérdésekhez (például Indiában a választásokon a félretájékoztatóról folytatott beszélgetések; Európában az MI és a magánélet miatti aggályok). Ez a beszélgetés lokalizációja azt jelentette, hogy a *Nexus* széles körű hatást gyakorolt, ösztönözve az információról és az MI-ről szóló beszélgetéseket sokféle kulturális kontextusban, nem csak az angol nyelvű médiában.

Források:

- Harari, Yuval Noah. *Nexus: Az információs hálózatok rövid története a kőkortól az MI-ig*. (Fern Press/Penguin Random House, 2024) – ahogy a Harari hivatalos oldalán és végzővegén is szerepel.
- Illing, Sean. "Yuval Noah Harari, hogy a demokrácia és az MI együtt létezhet-e." *Vox – The Gray Area* (Podcast interjú, 2024. szeptember 27.).
- Poole, Steven. "Nexus Yuval Noah Harari értékelése – a napok vége?" *The Guardian* (2024. szeptember 11.).
- Fox, Killian. "Kritika: Nexus: Az információs hálózatok rövid története..." *The Guardian/Observer* (2024).
- Smith-Ruiu, Justin. *Washington Post* kritikája a *Nexusról* (2024. október).
- Duncan, Dennis. "A 'Nexus'-ban Harari feltérképezi információs korunkat" *NYTimes Book Review* (2024. október).
- Green, Dominic. *Wall Street Journal* kritikája a *Nexus* című művéről (2024. szeptember).
- Immerwahr, Daniel. *Az Atlantic Nexus* kritikája (2024. október).
- Rubinstein, Samuel. *The Times (UK)* kritika (2024. szeptember).
- Könyvjelzők Nexus értékelések összesített gyűjteménye.
- Reddit vita témák a *Nexusról*.
- Barthold, Corbin. "Yuval Noah Harari fantasztikus forgatókönyvei" *Reason* (2024. szeptember 10.).
- Harari interjúk és előadások (Big Think, MSNBC, Wired, How To Academy stb.).



Kérlek, írd le Kate Crawford új könyvét, az "Atlas of AI"-t részletesen.

Az AI atlasza: átfogó elemzés

A fő érvek és tézis összefoglalása

Kate Crawford Atlas of AI (Yale University Press, 2021) átfogó kritikáját mutatja be a mesterséges intelligencia rejtett költségeinek és megerősített hatalmi struktúráinak feltárásával. Crawford központi tézise, hogy az MI "sem mesterséges, sem intelligens", hanem egy anyagi, kizsákmányoló ipar, amely összefonódik a meglévő társadalmi, politikai és gazdasági erőkhöz. Azt állítja, hogy az MI létrehozásához hatalmas erőforrás-hálózat szükséges – a földből bányászott ásványoktól kezdve az alacsony bérű munkáig, amely adatokat jegyzetel, egészen a felhasználók életéből gyűjtött személyes adatokig. Az MI rendszerek messze nem anyagtalannak vagy mágikusan autonómok lennének, hanem "embertelenítő extrakciókra" épülnek, amelyek kizsákmányolják a természeti erőforrásokat és az emberi munkaerőt. Crawford szerint ez a kitermelési alapú folyamat a technológiai vállalatok és kormányok kezében koncentrálja a hatalmat, ami antidemokratikus kormányzáshoz és mélyülő egyenlőtlenséghez vezet. Az MI Atlas nem kizárólag algoritmusok és kódok *szemszögéből tekintenek az AI-hoz, hanem anyagi és politikai nézőpontot képvisel, bemutatva, hogy az MI "kitermelési technológia", amely a társadalmat alakítja azáltal, hogy a kevesek érdekeit szolgálja a sokak rovására.*

Összefoglalva, Crawford fő érve, hogy az MI megtestesíti alkotóinak és támogatóinak – jellemzően nagyvállalatok és állami szereplők – értekeit, így gyakran felerősíti a meglévő igazságtalanságokat. A könyv szisztematikusan térképezi fel, hogyan szívja el az MI az erőforrásokat, kizsákmányolja a munkásokat, gyűjti fel az adatokat, és hogyan használja megfigyelési képességet, miközben objektívnek, elkerülhetetlennek és intelligensnek reklámozza. Azáltal, hogy lehúzzuk a szakápot ezen folyamatok előtt, *az Atlas of AI "értékes korrekcióként szolgál az MI körüli felhajtás nagy részére"*, arra ösztönözve minket, hogy az AI-t ne testetlen technikai csodaként, hanem *"anyagi, elfogult és saját nézeteinknek és ideológiáinknak alávetettként"* értsük. Ez az átfogó nézőpont megalapozza az MI ártalmainak visszafogását és annak fejlesztésének újragondolását a közérdek érdekében.

Fejezetről fejezetre történő elemzés

1. fejezet: *Föld* – Bolygói erőforrások kinyerése

Kulcsfontosságú meglátás: Az MI története a földben kezdődik. Az 1. fejezet a számításvány- és ökológiai költségeit vizsgálja. Crawford lítiumbányászati helyszínekre utazik (például Silver Peak, Nevadában), hogy bemutassa, hogyan támogassa az AI (adatközpontoktól eszközökig) a kitermelő iparágaktól függ, amelyek megsebzéseket okoznak a tájakat és hatalmas energiát fogyasztanak. Dokumentálja, hogyan bányászat a ritkaföldfémek és akkumulátorokhoz és elektronikához szükséges fémek kitermelése környezeti romláshoz, egészségügyi veszélyekhez a bányászok és elmenekült közösségek számára. Ez a fejezet alapfogalomként határozza meg a "számítási kitermelést": minden MI eszköz rejtett költséggel jár vízben, energiában és ásványokban. Crawford azt is megjegyzi, hogy hatalmas villamosenergia-fogyasztás szükséges a nagy AI modellek képzéséhez és szerverfarmok működtetéséhez. Az MI fizikai Földhöz való földjével a fejezet már az elejétől szembe kell néznünk az MI bolygóköltsegeivel.

2. fejezet: *Munka* – Emberi munkások az MI folyamatban

Kulcsfontosságú meglátás: Az "automatizált" MI rendszerek mögött egy emberi munkaerő hadserege áll. A 2. fejezetben Crawford azt a gyakran láthatatlan munkát vizsgálja, amely az MI-t működteti, a gyártól kezdve az adatkímkező asztaláig. Leírja az ember-robot hibrid munka jeleneteit az Amazon raktárakban és teljesítőközpontokban, összekapcsolva a mai MI-alapú munkahelyeket a munkaerő-optimalizálás hosszú történetével (a korai összeszerelő vonalaktól a gyorsétermi konyháig). A fejezet kiemeli az alacsony fizetésű, ismétlődő feladatokat és a munkavállalók megfigyelését: az alkalmazottakat a másodpercig követik, az MI rendszerek (például az Amazon algoritmikus vezetői) figyelik a teljesítményt és betartatják a termelési kvótákat. Crawford bevezeti a "fauxtomation" vagy "Potemkin MI" fogalmát, ahol az MI-vezérelt feladatokat valójában alulfizetett emberek végzik (például tartalommoderátorok vagy Amazon Mechanical Turk tömegdolgozók). Fontos példák az Amazon Mechanical Turk platformja, amely automatizált szolgáltatások illúzióját kelteti, miközben több ezer ember által végzett darabmunkát használja online. A fejezet tanulsága, hogy az MI alapvetően egy munkatörténet: ahogyan az MI ásványokat von ki, úgy emberi munkát *is kinyer*, gyakran kizsákmányoló körülmények között, hasonlóan a korai ipari munkavégzésekhez.

3. fejezet: *Adatok* – Az adatok értéke és politikája

Fontos meglátás: Az adatok az MI élettereje, de problémás módon gyűjtik össze. A 3. fejezet azt vizsgálja, hogyan táplálkoznak az MI rendszerek hatalmas adathalmazokról, amelyeket gyakran beleegyezés vagy magánélet tiszteletben tartása nélkül állítanak össze. Crawford nyomon követi az emberek adatainak gyűjtésének történetét – a büntetőfotóktól és korai megfigyelési képektől egészen a mai közösségi média tartalmak összegyűjtéséig –, hogy megmutassa: az "adatok" nem semleges forrás, hanem valami, amit emberektől és közösségektől vesznek el. Megjegyzik, hogy a korai időkben a kutatók AI adathalmazokat építettek ki kevés *felügyelet* nélkül: arcokat, hangokat és személyes adatokat szabadon bányásztak az internetből és a nyilvános nyilvántartásokból. Még ma is a technológiai cégek felhasználói adatokat (fotókat, bejegyzéseket, felvételeket) szívják fel az AI képzésének üzemanyagaként, gyakran jelentős beleegyezés nélkül – ezt a folyamatot Crawford a közösségi területek bezárásához hasonlítja, ahol a nyilvános adatokat magán MI-eszközökbe osztják ki. Például nagy arckép-adathalmazokról és ikonikus benchmarkokról, mint az ImageNet, tárgyalt, kiemelve, hogy ezek online képek kaparásával és az emberek címkézésével készültek, amelyek tükrözik a gyűjtők előítéleteit. Felidéri az Amazon sikertelen toborzási MI-jének figyelmeztető történetét is, amelyet korábbi felvételi adatokon képeztek, de végül elfogult lett a nők ellen –

ez egy példa arra, hogyan reprodukálhatja a hibás adatokon végzett képzés a diszkriminációt. Crawford tágabb érvelése az, hogy az adatok soha nem "nyers" vagy semlegesek – *"Minden adathalmaz... világnézetet tartalmaz"* – hangsúlyozza. Így az MI adatainak gyűjtése és címkézése alapvetően politikai gyakorlat, amely mély kérdéseket vet fel a magánéletéről, a tulajdonjogról és az elfogultságról.

4. fejezet: *Osztályozás* – A kategorizálás ereje

Kulcsfontosságú meglátás: Az MI megépítéséhez be kell sorolni a világot – ez egy hatalommal és előítéllettel teli cselekedet. A 4. fejezet azt vizsgálja, hogy az MI rendszerekben kiválasztott kategóriák és címkék hogyan tükrözik és erősítik meg az adott világnézeteket. Crawford szerint az, hogy az emberek vagy dolgok (nem, rassz, érzelem stb.) címkék eldöntése nem pusztán technikai lépés, hanem társadalmi hatalom formája. Történelmi párhuzamokat vont olyan gyakorlatokkal, mint a frenológia és a fiziognómia – olyan "tudományos" osztályozásokkal, amelyeket régóta hiteltelenítettek –, hogy megmutassa, milyen veszélyes elhinni az MI osztályozásainak objektívességét. Az ImageNetet esettanulmányként használva a fejezet feltárja, hogy még a széles körben használt MI-taknonómiák is képesek problémás feltételezéseket kódolni (az ImageNet, amely a WordNetre épült, örökölte a személykategóriák meghatározásában alkalmazott torzításokat). Crawford megjegyzi, hogy ami kimarad a besorolási rendszerből, ugyanolyan fontos, mint amit benne foglalnak: amikor az adathalmazok nem tartalmaznak bizonyos címkéket, vagy az adatokat durva kategóriákba kényszerítik, akkor gyakorlatilag eltörlik az identitásokat és árnyalatokat. Például az arcok bináris nemi vagy faj szerinti címkézése másokat tehet és félreértelmezheti az embereket, hasonlóan ahhoz, hogy az archaikus faji taxonómia tette. Egy feltűnő idézet a könyvből – *"Az osztályozások olyan technológiák, amelyek előállítják és korlátozzák a tudás módjait"* – hangsúlyozza, hogy minden osztályozási rendszer bizonyos értelmezéseket támogat, míg másokat elhallgattat. A fejezet azt is kiemeli, hogy ma sok MI modell saját fejlesztésű adathalmazokra (pl. Facebook vagy Google esetében) van kiképezve, ami azt jelenti, hogy a besorolási döntéseik fekete dobozokban zajlanak, amelyeket a nyilvánosság ellenőrzése elől védenek. Végző soron a 4. fejezet azt mutatja, hogy a látszólag hétköznapi tervezési döntések a címkékkel kapcsolatban átfogó etikai következményekkel járnak, mivel egy adott társadalmi víziót kódolnak be MI rendszerekbe.

5. fejezet: *Affektus* – Az érzelmefelismerés tévedése

Kulcsfontosságú meglátás: Az MI törekvése, hogy az emberi érzelmeket olvassa, kétséges tudományon alapul. Az 5. fejezetben Crawford az affektív számítástechnikához fordul – olyan algoritmusokhoz, amelyek azt állítják, hogy érzelmeket érzékelnek arcokból vagy hangokból. Kritikus történetet ír, különös tekintettel a befolyásos pszichológusra, Paul Ekmanra, akinek az univerzális arckifejezések elmélete számos érzelmefelismerő rendszer alapját képezi. Ekman Arc Akciókódolási Rendszere (FACS) az arcizommozgásokat állítólag univerzális érzelmekre kategorizálta, de Crawford feltárja, hogy Ekman kutatása színpadi kifejezéseken és szűk kulturális feltételezéseken alapult. A fejezet elmagyarázza, hogyan építettek olyan cégek, mint az Affectiva, mesterséges intelligencia modelleket ezekre a címkézett arc-adatbázisokra, és arra kérték a munkavállalókat, hogy érzelmekkel jelöljék meg az arcok képeit. Ez a megközelítés, Crawford szerint, továbbviszi Ekman alapötletének hibáit – túlzott leegyszerűsítést és félreértelmezést, mit jelent egy arckifejezés. Kiemeli, hogy az érzelmi kifejezések erősen kontextusfüggőek és kulturálisan változatosak, amit az egységes MI rendszerek nem vesznek figyelembe. Élénk példa erre az affektuszfelismerés alkalmazása az állásinterjúkon: a cégek kísérleteztek olyan MI-vel, amely elemzi a jelöltek arcvideóit, hogy személyiségüket vagy lelkesedésüket megítélje. Crawford figyelmeztet, hogy az ilyen gyakorlatok tele vannak

pontatlanságokkal és elfogultsággal – lényegében digitális frenológiával –, mégis bizonyíték nélkül alkalmazták őket, ami igazságtalanul alááshatja a jelentkezőket. Az affektív számítástechnika tudományos alapjait boncolva az 5. fejezet egy tágabb témát tár fel: amikor az MI azt állítja, hogy az emberi belső állapotokat (például érzelmeket) értelmezi, gyakran elavult vagy áltudományos elképzeléseket reprodukál, ami félreértés és megkülönböztetés kockázatát hordozza. Ez az érzelmi MI kritikája megerősíti a könyv üzenetét, miszerint *nem minden, amit mérni (vagy állítani mérhető lehet), objektívnek kell megbízni.*

6. fejezet: *Állam* – MI és a megfigyelés bővítése

Kulcsfontosságú meglátás: A modern MI párhuzamosan fejlődött a katonai és állami megfigyelési ambíciókkal. A 6. fejezet nyomon követi az MI és az állami hatalom eszközei közötti mély kapcsolatokat, a korai hidegháborús kutatásoktól egészen a mai rendőri algoritmusokig. Crawford megjegyzi, hogy az MI számos "innovációja" (mint például számítógépes látás, fordítás, autonóm járművek) eredetileg katonai finanszírozásból és védelmi projektekből származott. Ennek eredményeként az MI a hangsúlyt örökölt az osztályozásra, nyomon követésre és "ellenségek" célzására – olyan módszerekre, amelyeket ma már a civil életben is átalakítottak. A fejezet részletesen bemutatja, hogyan terjedtek el a hírszerző ügynökségeknél kezdődött fejlett megfigyelési eszközök, amelyek a helyi rendfenntartókig terjedtek. Például az arcfelismerő és előrejelző rendőrségi szoftvereket, amelyek korábban szigorúan titkosak voltak, most a rendőrség használja a városi utcákon. Crawford Edward Snowden leleplezéseire hivatkozik, leírva, hogyan teremtette meg a hírszerző ügynökségek "mindent összegyűjtő" adatkezelése (kommunikációs tömeges lehallgatás, metaadat-elemzés) a Big Tech adathalmazának mintáját. Beszél arról is, hogyan működnek együtt a technológiai cégek a kormánnyal: például a Project Maven, egy amerikai katonai MI-kezdeményezés, amely a Google mesterséges intelligenciáját használja drónfigyelő felvételek elemzésére. A Google rangos alkalmazottai megdöbbentek, amikor megtudták, hogy munkájukat militarizálják, ami belső tiltakozást váltott ki; A Google végül megszüntette a projektet, amely a védelmi vállalkozó, a Palantir céghez került. Egy másik eset a Cambridge Analytica botránya, amely bemutatja, hogyan használták az adatelemzést és az MI célzást politikai propaganda fegyvereiként. Ezek a példákon keresztül a 6. fejezet azt állítja, hogy az MI az állami hatalom és ellenőrzés kulcsfontosságú eszköze, gyakran elmosva a határt a nemzetbiztonság és a polgári megfigyelés között. Crawford szerint az MI politikája a dominancia politikája: akik az MI-t irányítják (kormányokat vagy technológiai óriásokat), túlzott képességeket szereznek a lakosság figyelésére, befolyásolására, sőt elnyomására. Ez megteremti a könyv végső hatalmi fókuszát.

Összegzés: *Erő* (7. fejezet) és *Coda: Tér* (8. fejezet)

Kulcsfontosságú meglátás: Az *Atlas of AI csúcspontja* megerősíti, hogy az MI alapvetően a hatalomról szól – kinek birtokolja és ki nincs-e –, és megkérdőjelezi a technológiai elit futurisztikus vízióit. A "Power" című Conclusion-ban Crawford összehangolja az előző fejezetek tanulságait, hogy azt állítsa, az MI megszilárdítja a meglévő hatalmi struktúrákat, hacsak nem avatkozunk aktívan. Azt állítja, hogy az "MI etikáról" beszélni nem elég; ami a *hatalomra való* fókuszra van szükség – annak megértése, hogy az MI *"arra van tervezve, hogy felerősítse és reprodukálja azokat a hatalomformákat, amelyekkel optimalizálni akarták"*. Más szóval, az MI rendszerek általában azok céljait szolgálják, akik megbízzák őket (nagyvállalatok, hadseregek, rendőrség), megerősítve dominanciájukat. A következtetés valószínűleg nagyobb demokratikus ellenőrzést követel az AI felett, és figyelmeztet attól, hogy az MI-t elkerülhetetlen, értéksemleges erőként fogadjuk el.

A következtetés után egy rövid Coda "Space" (8. fejezet) című koda egy provokatív végső határvonalra viszi az elemzést: a technológiai milliárdosok álmai az űr gyarmatosításáról. Crawford olyan alakokról beszél, mint Jeff Bezos és Elon Musk, akik elképzelik, hogy a Föld problémái elől úgy menekülnek, hogy pályára indulnak vagy más bolygókat telepítenek. Ezeket a ambíciókat a kitermelő kapitalizmus folytatásaként állítja be – egy új határok felfedezésére, amelyeket a Föld erőforrásai kimerítenek. Párhuzamokat vonva a gyarmati korszakokkal, azt sugallja, hogy a magánűrrepülés és a világon kívüli bányászat iránti törekvés szó szerint kiterjesztése az MI-ipari komplexum fenntarthatatlan fogyasztásának. Ebben a szucdában Crawford kritikája az MI "bolygóköltégeiről" zárul: ugyanaz a mentalitás, amely az MI-t arra ösztönzi, hogy kimerítse bolygónkat, most a kozmoszra összpontosít. Ez a befejezés egy kijózanító üzenetet erősít: irányváltás nélkül a technológiai ipar előrelépésre törekvése egyre nagyobb egyenlőtlenségekhez és környezeti pusztításhoz vezethet – akár a Földön és azon túl is. Ez egy végső felhívás, hogy újragondoljuk, mit jelent a fejlődés, és kik profitálnak belőle.

Az Atlas of MI főbb témái

Adatkivonás és zárás

A könyv egyik meghatározó témája az adat kinyerése. Crawford megmutatja, hogy az MI teljesítményei hatalmas mennyiségű adat felvételén alapulnak – gyakran az emberek tudtama nélkül kaparva vagy összeszedve. Ezt egy digitális földfoglaláshoz hasonlítja: a nyilvános és személyes adatokat természeti erőforrásként kezelik, amelyeket általában nagy technológiai cégek bányásznak. A beleegyezés és a magánélet etikai kérdései nagy jelentőséggel bírnak. Például emberek milliói képeit vették el a közösségi médiából vagy a webről, hogy a résztvevők engedélye nélkül készítsenek képzési adathalmazokat (például a ImageNet-et). Hasonlóképpen, a hangfelvételeket és az írott szövegeket összesítik, hogy virtuális asszisztenseket és nyelvi modelleket képezzenek, gyakran a felhasználók radarja alatt. Crawford hangsúlyozza, hogy ez a "beleegyezés vége" rendszer az emberi életet nyersanyaggá alakítja az MI számára, ami riasztást kelt a megfigyelés és az autonómia kapcsán. Továbbá az *Atlas of AI* azt állítja, hogy ha az adatokat MI rendszerek rögzítik és osztályozzák, azok gyakorlatilag saját platformokba kerülnek – még akkor is, ha az adatok eredetileg nyilvános voltak. A közös tevékenységünkől (posztok, képek, mozdulatok) kinyert érték néhány vállalati kézben van. Az adatbezárás témája rávilágít, hogyan privatizálhatja az MI a tudást ugyanúgy, ahogy az ipari kapitalizmus kerítette a földet (ezt Crawford kifejezetten a közös tulajdon és a bekerített terület eszméjével hozza fel). Röviden: az adatkitermelést nem áldozat nélküli "digitális kipufogó" gyűjtésként ábrázolják, hanem egy kizsákmányolási formájaként, amely gyakran elkerüli az informált beleegyezés normáit, és aránytalanul javát szolgál azoknak, akik képesek az adatok gyűjtésére és számítására.

Kizsákmányolt munkaerő az MI ökoszisztémában

Crawford az emberi munkát – amely gyakran láthatatlan a technológiai narratívákban – emeli a középpontba. Fő téma, hogy az MI kizsákmányolt munkaerőre épül, a bányáktól a mikromunkásokig. A könyv feltárja, hogy a kékgalléros és kattintásos munkák mennyire nélkülözhetetlenek az MI számára. A fizikai oldalon a bányák és hardverszerelő üzemek munkásai dolgoznak, hogy előállítsák az AI infrastruktúra alkatrészeit (néha veszélyes vagy alacsony bérű körülmények között). A digitális oldalon a *gig-dolgozók* címkéznek képeket, átírják a hangot, moderálják a tartalmakat, és elvégzik azokat a rutinfeladatokat, amelyek az MI rendszereket autonómnak látszik. Crawford részletezi az Amazon raktárakban és

tömegmunka platformokban zajló gyakorlatokat, hogy szemléltesse a kimerítő munkakörülményeket és a hipermegfigyelést: a dolgozókat algoritmusok figyelik sebesség és pontosság szempontjából, amelyeket a gép kiterjesztéseiként kezelnek. Azt is tárgyalja, hogy az automatizálás ígérete gyakran rejti el a *"Potemkin" megoldásokat* – például az AI chatbotokat, amelyeket titokban emberi munkások segítenek a nehéz lekérdezésekhez, vagy "automatizált" szolgáltatásokat, amelyek egy rejtett hadsereg támaszkodnak, akik a háttérben kattintanak és rendeznek. Ez a téma egy paradoxont emel alá: bár az MI-t gyakran úgy hirdetik, hogy megszüntetik az emberi munkaerőt, valójában átosszja a munkaerőt bizonytalanabb, alacsony fizetésű formákra (gyakran a globális délen vagy marginalizált közösségek között). Crawford munkaügyi fókusz felhívás arra, hogy felismerjük az MI ellátási láncában szereplő emberi költségeket és igazságtalanságokat. Arra ösztönzi a technológiai fejlesztőket és döntéshozókat, hogy fontolják meg a méltányos munkaügyi normákat – hogy az AI előnyei ne épüljenek a verejtékműhelyhez hasonló körülményekre vagy digitális munkakörön. A munka kiemelésével az *Atlas of MI* összhangban van azzal a növekvő megértéssel, hogy az MI etikának magában kell foglalnia a *munkavállalók jogait és a gazdasági igazságosságot*, nem csak a felhasználói magánéletet vagy az algoritmikus igazságosságot.

Az AI ökológiai és környezeti költségei

A környezeti hatás témája végigszövődik a könyvben, szó szerint az 1. fejezettől ("Föld") kezdve. Crawford feltárja, hogyan járul hozzá az MI fejlesztése a környezeti romláshoz és az éghajlatváltozáshoz, ami egy olyan szempont, amelyet a technológiai diskurzusban gyakran figyelmen kívül hagynak. A lítium és ritka fémek bányáitól és párologtató medencéitől kezdve az energiaigényes adatközpontokig és a mélytanulási modellek képzési folyamataiig az MI nagy szén- és anyaglábnymot mutat. A könyvben feltűnő példa egyetlen nagy MI modell képzésének összehasonlítása több autó élettartamának szén-dioxid-kibocsátásával – ami rámutat, hogy az "okos" algoritmusoknak lehetnek nagyon értelmetlen környezeti külső hatásai. Megjegyzi továbbá a rövid eszközök élettartama és az elektronikus hulladék fizikai hulladékait is. Ez az ökológiai költség témája összhangban áll a kitermelés gondolatával: az MI "kimeríti a bolygót" mind a véges erőforrások fogyasztásával, mind a fosszilis tüzelőanyagokból előállított áram fogyasztásával. Crawford elemzése szerint minden vidám technológiai bemutató vagy MI kütyü rejti el a *füstkérmények, olajkutak és szennyezett folyók láncát*. Ezeknek a hatásoknak a mennyiségi mérésével és az érintett környezetek történeteinek elmesélésével (például lítiumbányák vagy szerverfarmok körüli közösségek) az *Atlas of AI* arra ösztönzi az érintetteket, hogy vonják be a fenntarthatóságot az MI értékének számításaiba. Ez azt jelenti, hogy az MI fejlesztésének jövőjének olyan kérdésekkel kell szembesülnie, mint: Hogyan csökkenthetjük az AI energiafelhasználását? Újrahasznosíthatjuk vagy korlátozhatjuk a ritka ásványokat a hardverben? Ki viseli a "felhőket" (gyakran őslakos földeket vagy szegényebb régiókat) környezeti terhét? A könyv így a klíma- és környezeti igazságosságot az MI etika szerves részeként állítja be. A politikai szempontból ez a téma arra utal, hogy szükség van szabályozásokra vagy ösztönzőkre, hogy az MI fejlesztése környezetbarátabbá és elkötelezetesebb legyen az ökológiai lábnyoma miatt.

Politika és hatalmi struktúrák az MI-ben

Talán az *Atlas of MI legátható témája* az MI összefonódása a politikával és hatalommal. Crawford azt állítja, hogy az MI nem politikailag semleges – a társadalomban lévő hatalmi viszonyok alakítják és megerősítik. Ennek a témának egyik aspektusa a militarizáció és megfigyelés (amelyet a 6. fejezet "Állam" című fejezetben vizsgálunk), ahol látjuk, hogyan használják kormányzati szervek és befolyásos vállalatok az MI-t irányítási eszközként (a

tömeges megfigyelési programoktól a prediktív rendőrségig). Egy másik aspektus a gazdasági és vállalati hatalom: az MI fejlesztését ma néhány technológiai óriás (Google, Amazon, Microsoft, Facebook stb.) uralja, ami ezeknek az entitásoknak hatalmas befolyást biztosít az információra és az infrastruktúrára gyakorolni. Crawford rámutat, hogy az MI előnyei (és döntéshozatali joga) gyakran ezeknek a vállalatoknak és kormányzati partnereiknek szállnak meg, nem pedig a hétköznapi embereké. Például az arcfelismerés segíthet a rendőrségnek és növelheti a technológiai szolgáltatók profitját, de károsíthatja a közösségeket a félreazonosítással vagy a közösségi életre gyakorolt háborzongató hatásokkal. A könyv megmutatja, hogy az MI-t gyakran olyan módon alkalmazzák, amelyek erősítik a status quót – automatizálva az egyenlőtlenséget és kiterjesztve a hatalmasok elérését. Élénk példa erre az AI használata tartalommoderációban vagy felvételen: ezek az eszközök tükrözhetik alkotóik előítéleteit, kiszűrve bizonyos beszédeket vagy önéletrajzokat úgy, hogy tükrözzék a meglévő társadalmi előítéleteket (mégis "objektívként" jelennek meg). Crawford azt is tárgyalja, hogy a Big Tech ideológiai narratívái (például techno-solutionizmus, az az elképzelés, hogy az MI varázslatütésre megoldja a társadalmi problémákat) hogyan rejtik el a valódi hatalomrajongó helyzetet – ahol a közszektor egyre inkább a magántechológiára támaszkodik, és a polgárok adatgyűjtés és algoritmikus ítéletek tárgyává válnak. A könyv utolsó fejezetei ("Hatalom" és "Úr") kifejezetten megkérdőjelezzik a technológiai vezetők politikai vízióit, kritizálva, hogy libertárius vagy neokolonialista gondolkodásmódjuk (például a Marsra menekülni a Föld javítása helyett) elkerüli a demokratikus elszámoltathatóságot. Összességében ez a téma az MI újrapolitizálását követeli: *Ki döntheti el, hogyan építik és használják ezeket a rendszereket?* Crawford válasza figyelmeztetés, hogy jelenleg ezek a döntések nagyon kevés (főként nyugati, férfi, tehetős) kezekben vannak. Ez azt jelenti, hogy nagyobb közfelügyeletre, átláthatóságra és befogadásra van szükség az MI irányításában, hogy ellensúlyozzuk ezeket a hatalmi struktúrákat.

Az MI etikája és társadalmi következményei

Bár Crawford kritikusan kezeli a felszínes "MI etikai" beszédeket, könyve mélyen foglalkozik az AI alkalmazásának etikai és társadalmi következményeivel. Az elfogultság, az igazságosság és az MI emberi hatása témái végig megjelennek. Például közvetlenül foglalkozik az algoritmikus torzítással a tanítási adatok kontextusában (3–4. fejezet) és az affekt felismerés (5. fejezet). *Az Atlas of MI* konkrét példákat mutat be, amikor az MI rendszerek diszkriminatív eredményeket hoztak létre – például rasszista címkézést a képadathalmazokban vagy nemi torzítást a felvételi algoritmusokban –, hogy szemléltessen: az MI örökli a társadalom és alkotói előítéleteit. Egy másik etikai szempont az *AI képességeiről szóló hamis állítások*: Crawford megcáfolja azokat a mítoszokat, mint hogy az MI képes érzelmek vagy összetett tulajdonságok megbízhatóan érzékelni, ezeket nemcsak technikai hibáknak, hanem etikai hibáknak is nevezve, mert félrevezethetik és ártani az embereket (például egy "érzelmi MI" tévesen megbízhatatlannak jelöl egy állásjelöltet a rendszer etikai hibája). A könyv azt is vizsgálja, hogyan rombol a magánélet és a szabadság egy MI-vel átható világban. Összekapcsolja a mindenütt jelen lévő adatgyűjtést és a polgári szabadságjogokat fenyegető fenyegetést a pontszámot, ahogy az MI által működtetett megfigyelő rendszerek egyre általánossá válnak. Fontos, hogy Crawford nem korlátozza az etikai vitát az elvont elvekre; A gyakorlatokhoz és a hatalomhoz köti. Azt állítja, hogy nem elég az AI etikai irányelvek megalkotása, ha az alapvető üzleti modellek (adatkivonás, hirdetések, megfigyelés) változatlanok maradnak. Ehelyett a könyv azt javasolja, hogy vizsgáljuk, ki profitál és ki sérül minden MI alkalmazásban – lényegében igazságosság-orientált megközelítés. A társadalmi igazságosság témái kifejezetten előkerülnek: kiemeli, hogy az AI hatásait aránytalanul érzik a marginalizált csoportok (pl. gig-dolgozók, túlzott rendőrség alá szoruló közösségek, erőforrásokban gazdag, de szegény régiók

lakossága). Összefoglalva, az *Atlas of AI etikai szemüvege* az eredményekre és az erőegyensúlyhiányokra fókuszál. Crawford arra ösztönzi az olvasókat, hogy a "Ez igazságos-e ez az AI?" kérdéstől "Hogyan szolgálhatjuk az AI-t az egyenlőség és igazságosság szolgálatában?" kérdéséig – ami azt jelentheti, hogy bizonyos AI felhasználási módjait teljesen korlátozzák, ha azok ellentétesek az alapvető jogokkal. Az AI etika kiterjesztése a hatalom, munka és ökológia kérdéseire a könyv egyik legfőbb hozzájárulása.

Szerzői háttér és szakértelem

Kate Crawford szakmai háttere jelentősen erősíti az *Atlas of AI* tekintélyét. A mesterséges intelligencia társadalmi következményeinek vezető kutatója, két évtizedes karrierje a technológia társadalomra gyakorolt hatásainak vizsgálatában. Crawford magas rangú kutatói pozíciókat tölt be az akadémián és az iparban: kutatóprofesszor a USC Annenbergben, és hosszú ideig vezető kutató volt a Microsoft Researchnél, ahol társalapítója volt a Fairness, Accountability, Transparency, and Ethics (FATE) kutatócsoportnak. Társalapítója volt az AI Now Intézetnek az NYU-n, amely az egyik első kutatóintézet a mesterséges intelligencia társadalmi vonatkozásainak vizsgálatával foglalkozott. Ez az iparági és akadémiai tapasztalat ötvözi neki egy belső és külső nézőpontot – érti, hogyan fejlesztik az AI a vállalati laborokban, mégis kritikus, tudományos alapossággal közelíti meg.

Crawford szakértelme a könyv témáit is átöleli. Az USC-n munkája a technológiára fókuszál a történelem, a politika, a munka és a környezet kontextusában – lényegében ugyanazok a kvadránsok, amelyek az *Atlas of AI*-t is meghatározzák. Jelentős kutatásokat publikált olyan témákban, mint az adattorzítás, az AI etika és a hálózati megfigyelés vezető folyóiratokban (Nature, New Media & Society stb.). Emellett Crawford közös projektjei is kiemelik mesterséges intelligenciához való anyagi hozzáállását: például társalkotója volt az *"Anatomy of an AI System"* című díjnyertes vizuális térkép az Amazon Echo eszköz alkatrészeiről és ellátási láncáról. Ez a projekt szó szerint feltérképezte az ásványkitermelést, a munkaerőt és az adatáramlást, amely egyetlen "okos" előadót hajt – lényegében az *Atlas of AI* érveinek mikrokozmosza. Ez a munka azt mutatja, hogy Crawford évek óta kutatja az AI rejtett rétegeit, jóval azelőtt, hogy az már mainstream aggodalom lett volna.

A párizsi École Normale Supérieure-ben (a könyv írásának idején) első vendég elnökeként töltött be a mesterséges intelligencia és igazságszolgáltatás területén (a könyv írása idején) tovább jelzi az AI társadalmi igazságossági kérdéseire való fókuszát. Crawford tanácsot adott az ENSZ-, az Európai Parlament és más testületek döntéshozóinak is az AI-politikával kapcsolatban, jelezve, hogy tekintélyes tekintélyes az AI összetett hatásainak kormányzati keretrendszerekbe való átültetésében. Mindez a szakmai tapasztalat – tudósként, kutatóként, tanácsadóként – hozzájárul a könyv hitelességétől. Biztosítja az olvasókat arról, hogy az *AI atlasza* nem spekulatív polemikusan, hanem kiterjedt kutatás és területismeret terméke. Valóban, a könyv "több mint egy évtizednyi kutatásra" támaszkodik, beleértve helyszíni látogatásokat, interjúkat és interdiszciplináris tudományos munkákat. Crawford pozíciója a területen (gyakran említik más AI etikai kiemelkedő személyekkel együtt) azt jelenti, hogy az *Atlas of AI elemzése jelentős* jelentőséggel bír az akadémiai és szakpolitikai körökben. Ez a szakértelem lehetővé teszi számára, hogy összefonja a történelmet, a politikai gazdaságtan és a technikai ismereteket – összetett kapcsolatokat világossá téve, és hitelt adva az AI újragondolásának felhívásainak.

Kritikai fogadtatás és kritikák

Az *Atlas of AI* széles körű elismerést kapott tudósoktól, technológiai kommentátoroktól és a mainstream médiától, akik úttörő és időszerű MI kritikának tartották. A könyv gyorsan elengedhetetlen olvasmányként szerzett hírnevet: bekerült a Financial Times és a New Scientist év végi "2021 legjobb könyvei" listájára, és rangos elismeréseket nyert, mint például a 2022-es SHOT Sally Hacker Prize (kivételes technológiai kutatásért) és az ASIS&T Legjobb Információtudományi Könyvdíj. A kritikák túlnyomó többséggel dicsérték Crawford mesterséges intelligencia "munka és környezet kizsákmányolását", az "algoritmikus torzítást", valamint az MI képességeivel kapcsolatos "hamis állításokat" (például érzelmi felismerést). Anaïs Resseguier, az *AI and Ethics folyóiratban írva*, "alapvető munkának" nevezte a könyvet, és kiemelte annak jelentőségét, miközben a döntéshozók világszerte küzdenek az MI kárainak enyhítésével. A *Nature* című művében Virginia Dignum, az AI professzor dicsérte az *Atlas of AI*-t, amiért "feltárja az MI sikerének sötét oldalát", és "aprólékosan kutatott és kiválóan megírtnak" nevezte. A *Science* magazin kritikusa, Michael Spezio úgy jellemezte, hogy ez egy "átfogó kép a mesterséges intelligenciáról", amely azt mutatja, hogy az MI "gyorsan megszünteti a fenntartható jövő lehetőségeit globális szinten", hangsúlyozva, hogy Crawford hozzájárulása "időszerű és sürgős".

A média is megerősítette ezeket a pozitív értékeléseket. A *New Yorker* (egy "Briefly Noted" értékelésben) kiemelte Crawford tézisét, miszerint az MI "sem mesterséges, sem különösebben intelligens", lenyűgöző történetet kínálva a gépi tanulás alapját képező adatokról. A *New York Review of Books* jelentős kritikát közölt Sue Halperntől, aki hangsúlyozta a könyv könyörtelen feltárását az MI függöny mögötti "embertelenítő extraktív gyakorlatokról". Halpern idézete – amely Crawford honlapján szerepel – emlékezetesen összefoglalja a könyvet: "a mesterséges intelligencia nem deus ex machina formájában érkezik hozzánk, hanem számos embertelenítő, extrém gyakorlat révén, amelyekről a legtöbben nem vagyunk tudok.". A *Financial Times*-ban John Thornhill szerkesztő az *Atlas of AI*-t "értékes korrekciónak a [MI] hype-jára és hasznos útmutatónak a jövő számára", kiemelve, hogy Crawford erős keretet ad a technológiai forradalom veszélyeinek megértéséhez, és javasolja, hogyan irányíthatnánk ezt pozitív eredmények felé. David A. Shaywitz *Wall Street Journal* kritikája megjegyezte, hogy Crawford "szenvedélyesen érvel" az MI objektivitásának mítosza ellen, bizonyítva, hogy "míg az MI testetlennek [és] elkerülhetetlennek mutatják be, anyagi, elfogult és saját ideológiáinknak van alávetve.". Még olyan médiumok is, mint a *The Guardian* (John Naughton) és a technológiai kultúra oldalai is dicsérték a könyvet, mint "lenyűgöző" és alapvető olvasmányt az MI valós világbeli hatásairól.

A technológiai és az iparági hangok is üdvözölték a könyvet. Különösen Karen Hao, az *MIT Technology Review* akkori vezető szerkesztője azt írta, hogy Crawford élénk esettanulmányai az MI alapjairól "lehetetlenné teszik, hogy továbbra is pusztán absztrakt módon beszéljünk a technológiáról". Hao az *Atlas of AI*-t "mesterműnek" nevezte, mondván, hogy "nem tudta kiverni a fejéből". Ez az érzés beszédes, hiszen mélyen elmerült a technológiai tudósításban – ami arra utal, hogy a könyv még az MI szakemberek körében is megérintett, mivel szélesebb körű kontextust tár fel szakterületüknek.

Érdeemes megjegyezni, hogy bár a fogadtatás túlnyomórészt pozitív volt, néhány kritikus vitatta Crawford kompromisszumot nem engedő álláspontját. Például Michael Upshall egy tudományos publikációs blogon írt kritikája szerint az *Atlas of AI* "minden MI heves kritikája", ami azt sugallja, hogy Crawford az adott MI visszaéléseiről alkotott jogos kritikáit az egész terület vádjával vádolta. Upshall Crawford megközelítését egy modern luddita támadáshoz hasonlította, megkérdőjelezve, vajon túl széles ecsettel festett-e azzal, hogy minden MI rendszer eleve az elnyomó hatalomhoz igazodik. Rámutatott, hogy Crawford példái közül sok valódi

problémát emel ki – kizsákmányoló munkaerő, elfogult algoritmusok, megfigyelés –, de úgy érezte, nem ismerte el azokat az eseteket, amikor az MI-t etikailag vagy társadalmi jóból valószínűsítették meg. Ez a kisebbségi nézőpont rámutat egy feszültségre, ahogyan a különböző olvasók érzékelik a könyvet: az *Atlas of AI* nyíltan kritikus, és nem fordít sok időt az MI "fényes pontjaira". Crawford saját álláspontja, hogy néhány pozitív felhasználási területre való fókuszálás elvonhatja a figyelmet a szerkezeti kritikáról. Mégis, érdemes megemlíteni, hogy ez a kritika felmerült: azt mutatja, hogy Crawford könyve vitát vált ki. A bírálók túlnyomó többsége azonban egyetért abban, hogy még ha valaki nem is ítéli *el az összes* MI-t, ahogy néha látszik, az MI-ipar ellenőrizetlen gyakorlataival kapcsolatos problémái sürgetőek és valóságok.

Összefoglalva, az *Atlas of AI*-t mérföldkönek tartják a technológiai kutatások területén, és elismerést kapott tudományos folyóiratoktól, tech újságíróktól és kulturális kommentátoroktól egyaránt. A vizsgálóriadó, történelmi elemzés és etikai érvelés keverékét "*magisteriálisnak*" és "*nélkülözhetetlennek*" nevezték. Az MI teljes terepének feltérképezésével – a lítiumbányáktól a koncertmunkán át az adatszervereken át – Crawford átalakította az MI jövőjéről szóló beszélgetést. A fogadtatás azt mutatja, hogy a könyv nemcsak jól kutatott és jól megírt, hanem *nagy hatással van arra is, hogyan* gondolkodunk az MI társadalmi szerepéről, különösen az egyenlőség és igazságosság terén. Gyorsan kötelező olvasmánysá vált mindazok számára, akiket érdekel az MI politika, etika vagy a technológia és a társadalmi hatalmi struktúrák metszéspontja.

Főbb tanulságok és következmények

Crawford elemzése fényében az *Atlas of AI* több kulcsfontosságú tanulságot és következményt kínál a technológia, a szabályozási politika és a társadalmi igazságosság fejlődéséhez:

- Az MI egy társadalmi-technikai rendszer, amelynek valós költségei vannak: el kell adnunk azt az elképzelést, hogy az MI egy pusztán virtuális vagy "varázslatos" technológia. Minden MI modell vagy kütyü a természeti erőforrások, emberi munkaerő és adatok kitermeléséből származik. Következmények: A döntéshozóknak és mérnököknek figyelembe kell venniük a fenntarthatóságot és a munka etikát az MI rendszerek értékelésekor. Ez azt jelentheti, hogy iparági szabványokat kell meghatározni az MI képzés szénlábnyom-átláthatóságára, vagy igazságos béreket és munkakörülményeket biztosítani az "automatizált" szolgáltatások mögött álló emberek számára. Környezetvédelmi szabályozásokra lehet szükség az AI szén- és e-hulladék hatásának mérsékléséhez, és a munkaügyi törvényeknek kiterjedniük kellene a mecítés és a tömeges munkavállalókra is, amelyek elengedhetetlenek az MI fejlesztéséhez.
- Áttérés az "etikai mosásról" a hatalomra és elszámoltathatóságra: Crawford alapvető érve, hogy nem elegendő magas szintű AI etikai elveket kiadni, miközben figyelmen kívül hagyjuk az AI építésének és alkalmazásának hatalmi egyensúlyhiányait. Következmény: Az MI körüli diskurzus a konkrét elszámoltathatósági intézkedések és hatalmi ellenőrzések felé kellene mozdulni. Ez magában foglalhatja a Big Tech (versenyellenes intézkedések, adatirányítási törvények) nagyobb szabályozási felügyeletét, hogy megakadályozza az MI erő koncentrációját. Ez azt is jelenti, hogy sokszínű érintetteket – különösen azokat a közösségeket, amelyeket gyakran az MI károsít (munkavállalók, marginalizált csoportok stb.) – bevonni a fejlesztési és irányítási folyamatba. Ahelyett, hogy csak azt kérdeznénk: "Ez az algoritmus igazságos-e?", inkább azt kell kérdeznünk, hogy "Kit szolgál ez az MI, és kinek a költségére?", és ennek megfelelően kell megterveznünk a beavatkozásokat.

- Adatjogok és magánéletvédelem: Az adatok kinyeréséről szóló széles körű vitát figyelembe véve a legfontosabb tanulság, hogy az egyéneknek és közösségeknek nagyobb kontrollra van szükségük adataik felett. Következmények: A szabályozók megerősíthetnék az adatvédelmi törvényeket (például GDPR eszközökre építve), hogy megköveteljék az MI képzésben használt adatok tájékozott beleegyezését és átláthatóságát. Előfordulhat új adatkezelési formák is, mint például adatbizalmi bizások vagy kollektív tárgyalások (hogy az emberek ne csak adatalányok legyenek, hanem beleszólásuk legyen a kollektív adataik felhasználásába). A könyvben kiemelt "beleegyezés vége" kezelése bizonyos adatgyűjtési vagy -értékesítési formák betiltását jelentheti, különösen érzékeny területeken, mint az arcfelismerés esetében.
- Munkaügyi reformok az MI gazdaságban: *Az Atlas of AI* egyértelművé teszi, hogy az automatizálás mögött sok alulfizetett dolgozó áll. Következmények: Ahogy az MI terjed, a munkaügyi politikáknak alkalmazkodniuk kell ahhoz, hogy megvédjék azokat, akik az MI árnyékában dolgoznak. Ez magában foglalhatja a méltányos bérgaranciákat a tömegmunkások számára, átláthatóságot a vállalatoktól arról, mennyi emberi munkaerő van az "MI"-jükben (hogy megelőzzék a megtévesztő marketinget és megfelelően értékeljék a munkát), valamint támogatni azokat a munkavállalókat, akiket az MI-vezérelt vezetés elhelyezett vagy megfigyelt. Tágabb szinten a társadalomnak meg kellene kérdőjeleznie az elkerülhetetlen automatizálás narratíváját, és inkább azt kell terveznie, hogyan lehet olyan módon használni, hogy az emberi munkahelyeket javítsa anélkül, hogy kizsákmányolná az embereket – például a munkavállalók bővítésével és a szakképzett felügyelet értékelésével, ahelyett, hogy a munkavállalókat közvetlenül helyettesítenénk vagy "MI" álarc mögé rejtjük őket.
- Az egyenlőség és a társadalmi igazságosság központi célok: Crawford elfogultság, megfigyelés és hatalom témái a polgári jogok és társadalmi igazságosság kérdéseivel kapcsolatos kérdéseket érintik. Következmények: Bármely MI irányítási keretrendszernek alapvető elvként kell kezelnie a méltányosságot. Például olyan területeken, mint a büntető igazságszolgáltatás, lakhatás, egészségügy és oktatás, az MI használatát ellenőrizni kellene a faji kisebbségekre, szegényekre vagy más sebezhető csoportokra gyakorolt eltérő hatások szempontjából. Az olyan eszközöket, mint a prediktív rendőrség vagy az automatizált felvétel, szigorú ellenőrzésnek vagy moratóriumnak kellene alávetni, ha megállapítják, hogy megerősítik a diszkriminációt. A közösségi hozzájárulás és felügyelet (például MI-felülvizsgálati bizottságok, amelyekben állampolgárok is vannak) segíthet abban, hogy az AI bevezetése összhangban legyen a közértékekkel. Lényegében az MI társadalomba való integrációját proaktív intézkedéseknek kell kísérnie, hogy megakadályozzák az egyenlőtlenség felerősödését, amire Crawford figyelmeztet.
- Techno-utópikus narratívák újragondolása: A könyv végső kritikája a milliárdos űrfantáziákról és az átfogó vállalati narratívákról azt sugallja, hogy ki kell kérdőjeleznünk a technológia önmagáról elmesélt domináns történeteit. Következmény: Ahelyett, hogy elfogadjunk azt a retorikát, hogy az MI elkerülhetetlenül megoldja a nagy problémákat, vagy hogy a "haladás" azt jelenti, bármit is követnek a Szilícium-völgyi milliárdosok, demokratikus mérlegelésnek kellene lennie az innováció irányáról. A társadalom követelheti, hogy az innováció a valódi emberi igényeket (mint például klímaalkalmazkodás, egészségügy, igazságos oktatás) *itt a Földön* kielégítse, nem pedig túlzó projekteket, amelyek az elit érdekeit szolgálják. Ez azt is jelenti, hogy támogatják a kutatási és fejlesztési modelleket, amelyek nyitottak, elszámoltatottak és közjót céloznak – például az MI kutatás finanszírozását egyetemeken vagy nonprofit szervezetekben, amelyek társadalmi problémákra fókuszálnak, ahelyett, hogy a vállalatok kezére bíznák, amelyek a profit vagy katonai célok vezérelnek.

- "Visszautasítás" és Kollektív Cselekvés: Crawford egy "visszautasítás" felhívásával zárul – közösségek és társadalmak, akik bizonyos káros mesterséges intelligencia felhasználását elutasítják, és alternatív utakat követelnek. Következménye: Előfordulhat, hogy a legjobb megoldás az MI alkalmazásának elkerülése. Például a városok elutasíthatják az arcfelismerést a rendőrségben annak polgári szabadságjogokkal kapcsolatos költségei miatt; az iskolák elutasíthatják az invazív felügyelő MI-t, amely gyanúsítottként kezeli a diákokat. Ez a koncepció felhatalmazza a közvéleményt arra, hogy nemet mondjon azoknak a technológiáknak, amelyek ellentétesek az értékeikkel. Ez összhangban áll azzal az elképzeléssel, hogy a társadalmi mozgalmak (a munkavállalói jogokért, adatigazságosságért, környezetvédelemért) kulcsfontosságúak lesznek egy igazságos MI jövő kialakításában. Crawford hangsúlyozza, hogy értelmes változás "csak akkor történhet meg, ha megkérdőjelezzük azokat a hatalmi struktúrákat, amelyeket az MI jelenleg megerősít, és megteremtjük egy másik társadalom alapjait." Ez egy közös cselekvésre való felhívás: a technológusok, állampolgárok, aktivisták és törvényhozók együtt dolgoznak azon, hogy az MI fejlesztését az igazságosság, fenntarthatóság és a közérdek felé irányítsák.

Összefoglalva, az *Atlas of AI* nemcsak a mesterséges intelligencia jelenlegi pályájának problémáit diagnosztizálja, hanem egyben világos felhívásként is szolgál arra, hogy újragondoljuk, hogyan tervezzük és irányítjuk ezeket a technológiákat. A könyv meglátásai arra ösztönöznek minket, hogy az MI *a mi feltételeink* szerint haladjon – az emberi méltóság, az egyenlőség és a bolygó jóléte előtérbe kerülve – ne pedig néhány hatalmasabb ember feltételei szerint. Ezeknek a tanulságoknak a belsővé tételével a technológia és a politika területén dolgozók elkezdhetnek olyan MI rendszereket építeni, amelyek valóban szolgálják a társadalmat, és olyan szabályozásokat és gyakorlatokat vezethetnek be, amelyek megakadályozzák az AI előnyei és terhei koncentrációját olyan módon, ahogy Kate Crawford oly meggyőzően figyelmeztetett.

Források:

1. Kate Crawford, *Az MI atlasza: Hatalom, politika és a mesterséges intelligencia bolygói költségei*. Yale University Press, 2021 – Kiadó leírása és szerzői életrajza.
2. *Az MI Atlasz áttekintése* – Wikipédia.
3. Francesca Rossi, *Atlas of AI áttekintése, Artificial Intelligence folyóirat (2022)* – fejezetösszefoglalók.
4. Mindy McAdams, "Könyvjegyzetek: Atlas of MI" – AI a médiában és társadalomban blog (2021. augusztus).
5. Michael Upshall, "Az MI atlasza" kritikája – Thinking About Digital Publishing (2021. december).
6. Kate Crawford honlapja – *Atlas of MI* oldal és róla oldal (szerzői képesítések, szerkesztőségi kritikák).
7. Yale University Press – *Atlas of AI* könyvoldal (sajtókritika és elismerések)

Kérjük, részletesen írd le Kai-Fu Lee és Chen Qiufan új könyvét, a "Artificial Intelligence 2041"-et.

Kai-Fu Lee és Chen Qiufan "AI-2041: Tíz vízió a jövőnkre"

Ez egy ambiciózus könyv, amelyet Kai-Fu Lee – tapasztalt MI-szakértő és volt Google China elnök – valamint Chen Qiufan (Stanley Chen) – díjnyertes kínai sci-fi író társszerzői. Együtt tíz futurisztikus forgatókönyvet mutatnak be 2041-re, mindegyik egy lebilincselő novellát egy ismeretterjesztő esszével párosítva. Ez a hibrid formátum lehetővé teszi a szerzők számára, hogy történetmesélésen keresztül ábraszítsák a hihető, mesterséges intelligencia által vezérelt jövőket, majd elemezzék a történetek mögött álló valódi technológiákat és trendeket. Az eredmény egyszerre élénk narratív út és egy "gyorsút" az MI társadalomra gyakorolt potenciális hatásáról. Az alábbiakban részletes áttekintés található a tíz vízióból, a kulstechnológiákról és azok előrevetített fejlődéséről, az átfogó témákról (társadalmi, gazdasági, etikai), a könyv formátumáról és céljáról, a szerzők profiljáról, valamint arról, hogyan *fogadták az AI 2041* a szakértők és a közönség körében.

Szerzők és koncepcionális megközelítés

Kai-Fu Lee évtizedes mesterséges intelligencia szakértelmet hoz a projekthez. Ph.D. számítástechnikus (Carnegie Mellon) és technológiai vezető, vezette az AI kutatásait az Apple-nél és a Microsoftnál, megalapította a Google China-t, és jelenleg a Sinovation Ventures céget vezeti. Lee korábbi könyve, *az AI Superpowers* kiemelte az AI fejlesztéseket Kínában és az Egyesült Államokban, és az *AI 2041-ben* ezt a előrelátást a következő 20 évre is kiterjeszti. Társelnöke a World Economic Forum AI Council-nek, és ismert arról, hogy keretezi az MI irányvonalát (például a "Négy Hullám MI" keretrendszerét), valamint a proaktív alkalmazkodás (például a munkahelyi újraképzés és *az ember-mesterséges intelligencia szimbiózis*) támogatását. Az *AI 2041-ben* Lee minden történet után elemző esszéket ír, elmagyarázva, hogyan működnek a jelenlegi MI technológiák (pl. mélytanulás, NLP, robotika) ma és hogyan fognak fejlődni 2041-re, valamint hogyan dolgozik ki a politikai és üzleti betekintést ezeknek a változásoknak a kezelésére. Figyelemre méltó, hogy Lee a spekulációkat a "realista MI"-re korlátozza – olyan fejlesztéseket, amelyeknek becslése *szerint* >80%-os valószínűséggel valósulnak meg 2041-re –, elkerülve a távoli jövőbeli kliséket, mint az érző robotokat. Ahogy Lee fogalmaz: *"még kevés vagy semmilyen áttörés ellenére is az MI továbbra is készen áll arra, hogy mélyreható hatást gyakoroljon társadalmunkra."*

Chen Qiufan viszont elismert fikciós író, aki a "sci-fi realizmusáról" ismert. Olyan regények szerzője, mint a *Waste Tide*, Chen gyakran ötvözi a futurisztikus technológiát társadalmi kommentárral (Kína technológiai jövőjének jósnőjének nevezték). Az *AI 2041-ben* Chen tíz novellát ír, amelyek a világ különböző pontjain játszódnak (Mumbai, Lagos, Sanghaj, Tokió stb.), mindegyik az MI mindennapi életre gyakorolt hatásának egy egyedi aspektusát mutatja be. Ezek a történetek nem disztópikus fantáziák; valós technológiákon és hihető trendeken alapulnak, amelyeket mindössze két évtizeddel előre előre extrapolálnak. Chen képzeletgazdag narratívái emberi arcot vetnek az absztrakt technológiának: diákokkal, művészekkel, hackerekkel, orvosokkal, sofőrökkel és másokkal, akik etikai dilemmákkal és személyes kihívásokkal szembesülnek az AI-gazdag jövőben. Chen történetmesélését Lee magyarázataival ötvözi a könyv azt reméli, hogy "emberibbá teszi az MI lehetséges kárait és előnyeit", és felébreszti az olvasók képzeletét az MI-vezérelt világunkról. A szerzők kifejezetten azt akarták bemutatni, hogy egy olyan jövőt *ábrázoljanak, amelyben élni szeretnének és alakítani szeretnének, megerősítve "hitüket az emberi cselekvésben – hogy mi vagyunk sorsunk urai, és egyetlen technológiai forradalom sem fogja ezt megváltoztatni"*. Ez az optimista szellemiség sok történetet alapoz, még akkor is, ha komoly kockázatokat is elismernek.

Formátum: Fikció + elemzés minden vízióhoz

A könyv szerkezete minden fejezetben váltakozik a fikció és a nem-fikció között. Általában egy fejezet Chen novellájával kezdődik, amely egy 2041-es világot dramatizál, amelyet egy adott MI technológia alakít át. A történet elmerül egy olyan narratívában, amely kiemeli a technológia ígérteit és buktatóit – gyakran a főszereplő személyes küzdelmén vagy erkölcsi döntésén keresztül. A történet után Kai-Fu Lee esszéje a bemutatott technológiát bontja: elmagyarázza a jelenlegi csúcstechnológiát, a legutóbbi áttöréseket és a várhatóan következő 20 év trendjeit. Elérhető értelemmel tárgyalja a technikai fogalmakat (*a konvolúciós neurális hálózatoktól a kvantumszámításig*), és feltárja a történet által felvetett társadalmi következményeket (pl. elfogultság, magánélet, foglalkoztatás, biztonság). Ezt a hívás-válasz formátumot "zseniálisnak" és nagyon hozzáférhetőnek dicsérték – gyakorlatilag egy "szemfélnyitó" esettanulmányok sorozata, amely az AI absztrakt problémáit konkrétvá teszi. Így minden fejezet "kapuként" működik az MI jövőnk egyik aspektusába. A végére a tíz vízió együtt széles körű területeket fed le: oktatás, egészségügy, pénzügy, szórakoztatás, közlekedés, hadviselés, foglalkoztatás, kormányzás, magánélet és közgazdaságtan.

Az alábbiakban az összes tíz fejezet (vagy "látomás") összefoglalója található, beleértve minden történet összefoglalóját, valamint a kapcsolódó technológiákat és betekintéseket:

A Tíz Látomás – fejezetről fejezetre áttekintés

1. "Az aranyelefánt" – AI biztosítás és algoritmikus torzítás (Mumbai, 2041)

Az első történetben egy indiai tinédzser, Nayana 2041-ben él Mumbaiban a *Ganesh Insurance* felügyelete alatt, amely egy mélytanulásra épülő program, amely családjá minden aspektusát nyomon követi. A család örömmel cseréli a magánéletet drámaian alacsonyabb biztosítási díjakért: telepítik a biztosító alkalmazásait vásárlásra, vezetésre, étrendre, testmozgásra – sőt hidratálásra is –, és folyamatos ösztönzéseket kapnak, hogy "optimálisan" éljenek. Eleinte az AI ajánlásai ártalmatlannak és hasznosnak tűnnek (Nayana apja még a dohányzásról is leáll a kitartó utasítások miatt). Azonban amikor Nayana barátságot köt egy Sahej nevű fiúval a

virtuális iskolájában, egy nyugtalanító mellékhatás merül fel: az algoritmus hirtelen emeli a család díjait. Kiderült, hogy Sahej egy alacsonyabb kasztból származik (dalit), és bár az MI sosem volt kifejezetten kasztbeli torzításra programozva, mély neurális hálózata érzékelt némi összefüggést észlelt a társas kapcsolatok és a kockázat között. A biztosítási kockázati profilok optimalizálásakor az MI gyakorlatilag "megtanult" egy diszkriminatív előítéletet – megbüntetve Nayanát, amiért egy történelmileg marginalizált kasztból származó fiúval társult meg. Nayana anyja, aki aggódik a költségek miatt, nyomást gyakorol rá, hogy kerülje Sahejet, így Nayanát arra kényszeríti, hogy az algoritmikus irányítás és a személyes lelkiismeret között válasszon. Végül lázad, és érvényesíti emberi önállóságát, nem engedi, hogy egy fekete doboz modell irányítsa kapcsolatait.

Lee *"Az aranyelefánt"* elemzése bemutatja az MI külső hatásainak fogalmát – olyan nem szándékolt káros következményeket, amelyek akkor merülnek fel, amikor egy MI egyirányúan optimalizál egy szűk célt. Ahogy ma a közösségi média algoritmusok maximalizálják az elköteleződést, de felerősítik a felháborodást és az elfogultságot, itt egy egészségoptimalizáló MI végül beágyazza a társadalmi egyenlőtlenséget. A fejezet mélytanulás működését vizsgálja: hatalmas adathalmazokon (jelen esetben a felhasználók egészsége, pénzügyei és társadalmi adatai) történő képzéssel a többrétegű ideghálózatok olyan mintákat fedezhetnek fel, amelyeket az ember nem veszne észre. De ha az adatok történelmi torzításokat tükröznek, az MI *matematikailag* reprodukálja ezeket az elfogultságokat, miközben "objektívnek" tűnik. Lee elmagyarázza, hogy 2041-re az ilyen "életoptimalizáló" MI szolgáltatások elterjedhetnek – az MI valóban *"jobban ismerheti a felhasználókat, mint ők magukat"* és folyamatosan befolyásolhatja a viselkedést. Ez előnyöket kínál (egészségesebb szokások, megelőző ellátás), de figyelmeztető jeleket is felvet: a magánélet eróziója, a beleegyezés hiánya, valamint a "társadalmi hitel"-szerű rendszerek kockázata, ahol a lehetőségek (vagy biztosítási díjak) átláthatatlan MI-ítéletektől függenek. A legfontosabb tanulság az, hogy az MI nem *alapvetően erkölcsös* – ha csak a profit vagy hatékonyság maximalizálására képezzük, akkor az ősi előítéleteket fenntarthatja a modern algoritmusokkal. Ez a nyitó történet tehát egy olyan témát teremt, amely végigvisszhangzik a könyvben: *Hogyan lehetne az MI-t az emberi értékekhez és a méltányossághoz igazítani?*

(Elemzési témák: mélytanulás, big data, az AI biztosítási/pénzügyi alkalmazásai, algoritmikus torzítás/külső tényezők.)

2. "Az istenek a maszkok mögött" – Deepfakes és igazság az MI korában (Lagos, 2041)

A második látomás Lagosba, Nigériába visz minket, ahol az általános megfigyelő kamerák és az arcfelismerés egy virágzó technológiai szcénával együtt élnek 2041-ben. A történet Amaka nyomában áll, egy tehetséges fiatal programozó, aki deepfake-eket – hiperrealisztikus MI-generált videókat – készít. Amakát egy igbo etnikai közösséget képviselő csoport csábítja be egy titkos munkára, akik az ő javukra akarnak fordítani egy országos választást. Rákényszerítik, hogy hamisítson egy videót egy rivális politikusról, aki botrányt követ el, és azzal fenyegetőznek, hogy zsarolják Amakát saját deepfake eszközükkel – egy olyan videóval, amely kompromittáló helyzetben ábrázolja, amely a konzervatív nigériai társadalomban tönkretethetné az életét. Ez a baljós történet kiemeli az MI kétélű erejét a médiában – 2041-re a GAN által generált ("generatív ellenséges hálózat") videóhamisítások annyira hibátlanok, hogy a közvélemény (vagy akár a rendőrség) számára szinte lehetetlen megmondani, mi a valóság. Lagosban a fiatalok nemcsak divatként viselnek stilizált maszkokat, hanem az arcfelismerő kamerák megfélemezésére is, és maga az igazság vitatott áruvá vált. A cím "maszkok mögötti

istenek" kifejezése arra utal, hogy rejtett szereplők mesterséges intelligenciát használnak, hogy bábmestert játszanak a társadalom felfogásával.

Lee kommentárja ehhez a fejezethez a számítógépes látás és a generatív modellezés fejlődését vizsgálja. Elmagyarázza, hogyan működnek a GAN-ok: egy "generátor" hálózat hamis képeket vagy videókat gyárt, míg egy "diszkriminátor" hálózat próbálja észlelni a hamisítványokat – a két mesterséges intelligencia verseng, és fejlődik, míg a hamisítványok megkülönböztethetlenné válnak a valóságtól. 2021-re a GAN-ok már nagyon meggyőző hamis arcokat tudnak előállítani; A könyv szerint 2041-re a deepfake technológia annyira fejlett és elterjedt lesz, hogy komoly biztonsági fenyegetést jelent. Lee a következményeket tárgyalja: a manipulált videók erőszakot szíthatnak, hírnevet tönkretelhetnek, vagy alááshatják a demokráciát azzal, hogy elárasztják az infoszférát dezinformációval. Ez a forgatókönyv *"a ma tanúi turbo feltöltött változata"* – jegyzi meg a könyv. Az elemzés érinti az AI biztonsági intézkedéseket is – például digitális vízjeleket vagy deepfake észlelési algoritmusokat –, de figyelmeztet a macska-egér dinamikára is, ahogy a hamisítványok javulnak. A tágabb téma a bizalom eróziója: amikor a látás már nem hisz hiszélés, a társadalmak káoszba vagy autoriter megszorításokhoz fordulhatnak. Az információk hitelességének biztosítása kiemelten fontossá válik. Összességében a "Gods Behind the Masks" élénken szemlélteti az MI etikáját a médiában: vajon az MI képes lenne a kreatív kifejezést és reprezentációt, vagy *példátlan* léptékben a megtévesztést?

(Elemzési témák: számítógépes látás, konvolúciós neurális hálózatok, mélyfeszítés GAN-okon keresztül, biometria és megfigyelés, mesterséges intelligencia kiberbiztonság.)

3. "Twin Sparrows" – Személyre szabott AI oktatók és az oktatás jövője (Szöul, 2041)

Ez a fejezet azt mutatja be, hogyan alakíthatják át a természetes nyelvfeldolgozás (NLP) és a személyre szabott mesterséges intelligencia társai a tanulást. A történetben, amely Dél-Koreában játszódik, az ikertestvérek (Jihee és Jihan) árvák lesznek, és egy modern árvaházba kerülnek, amely mesterséges intelligencia gondozókkal van felszerelve. Minden ikerhez egy AI oktató tartozik – egy interaktív, rajzfilmszerű karakter (saját "veréb"), aki kiterjesztett valóságban él és velük együtt fejlődik. Ezek az AI mentorok fejlett NLP-t használnak (amely a GPT-3 leszármazottaira emlékeztet) természetes beszélgetéshez, mélyen egyénileg tanítva a gyerekeket. Ahogy az ikrek haladnak az iskolában, az MI oktatók dinamikusan alkalmazkodnak személyiségükhöz, tanulási stílusukhoz és érzelmi szükségleteihez. Az egyik iker félénk és művészi, a másik nyitott és elemző; AI társaik ennek megfelelően alakítják ki az órákat és kihívásokat, így egyben helyettesítő tanárokká, tanácsadókká és barátokká válnak. A házi feladat játékos párbeszédmé válik, és az osztálytermi órákat személyre szabott AR átfedéseken keresztül adják át – így minden diák megláthatja és hallhatja egy nekik optimalizált lecke változatát. A történet kiemeli az előnyöket (egyetlen gyerek sem csúszik ki a rések közül; az MI türelmesen kezeli minden diák gyengeségeit és motiválja őket), valamint a lehetséges mellékhatásokat. Az ikrek intenzív érzelmi köteleket alakítanak ki AI verébeikkel – felveti a kérdést, hogyan illeszkednek az emberi tanárok vagy kortársak egy olyan világba, ahol egy gyerek azt mondhatja: "az AI barátom jobban ért, mint bárki más." Valóban, a következő generáció ebben a világban az MI társakat *"önmaguk természetes kiterjesztéseinek"* látja, talán még intimebbnek, mint az emberi kapcsolatoknak. Ez arra utal, hogy túlzott függőség vagy akár függőség merül fel a tökéletesen megértő MI-től (a mai okostelefon csatlományból kiindulva).

Az elemzés során Lee a természetes nyelvi modelleket és az MI-t vizsgálja az oktatásban. Áttekinti az ugrást a korai NLP-től a hatalmas transzformátor-alapú modellekig, mint például a GPT-3 (amely 2020-ban emberi, szöveggenerálásával megdöbbsentette a megfigyelőket). 2041-re egy olyan GPT-23-at képzel el, amely gyakorlatilag minden ember által létrehozott tartalmat elolvasott és megnézte, lehetővé téve gyakorlatilag bármilyen kérdésre válaszolni vagy mély beszélgetéseket folytatni. Ez a fejlődés rendkívül kompetens AI oktatókat tehet – képesek bármilyen tantárgyat, bármilyen nyelven vagy stílust tanítani, és ezt egyénileg is megtehetnék. A fejezet megjegyzi, hogy az MI már most is mutatja ennek a pillanatait: például a GPT-3 utódjai és beszélgető ügynökei (például a ChatGPT, amely egy évvel a könyv után jelent meg) bemutatják, hogyan tud egy MI személyre szabott segítséget vagy kódolási tanácsot adni. Lee szerint ezek az MI-oktatók demokratizálhatják az oktatást, elhozva a felső szintű oktatást vidéki falvakba vagy alulfinanszírozott iskolákba, és lehetővé tennék, hogy minden diák a saját tempójában tanuljon. Ugyanakkor kritikus kérdéseket is felvet: *Mi az emberi tanárok szerepe?* Talán mentorokká válnak, akik a szocio-érzelmi fejlődésre fókuszálnak, miközben az MI a rutinos tanítást végzi. És *mi a helyzet a pszichológiai hatással?* Ha a gyerekek kötődnek az AI barátokhoz, figyelembe kell vennünk az empátiát és a társas készségeket érintő következményeket. Az elemzés érinti az önfelügyelő tanulás fogalmát (ahogyan a modellek magukat tanulnak címkezetlen adatokból, emberiesebb megértéssel), sőt az MI tudatosságának lehetőségét is, bár Lee szkeptikus, hogy az AGI valóban 2041-re fog megjelenni. Végző soron a "Twin Sparrows" optimista, mégis óvatos szemléletet nyújt az MI oktatásban – megmutatja, hogyan ápolhatja minden gyermek potenciálját, miközben figyelmeztet, hogy egy olyan MI, amely "egyedülállóan ért" az emberi kapcsolatokat kiszámíthatatlan módon megváltoztathatja.

(Elemzési témák: természetes nyelvfeldolgozás, GPT-3 és annál tovább, önfelügyelő képzés, mesterséges intelligencia a személyre szabott oktatásban, AGI és tudatosság megvitatása.)

4. "Érintés nélküli szeretet" – MI a járdémia utáni világban: egészségügy és virtuális kapcsolatok (Sanghaj és São Paulo, 2041)

A történet egy elhúzódó COVID-szerű járvány utóhatásai fonaljában vizsgálja, hogyan teheti lehetővé az MI, a biotechnológia és a robotika az életet távolról. A "Kontakt nélküli szerelem" című részben egy fiatal pár – egy Kínában, egy Brazíliában – mély, távolsági kapcsolatot tart fenn egy teljesen magával ragadó virtuális világon keresztül. A fizikai utazás továbbra is korlátozott a folyamatos vírusvariánsok miatt, de a technológia alkalmazkodott: az emberek fejlett "bőrölelésű" öltönyöket és AR szemüveget viselnek, hogy közös virtuális terekben találkozzanak, és a kontinensek távolsága ellenére is szimulált érintést éreznek. Eközben a személyes egészség minden aspektusa digitalizált. A karaktereknek viselhető vagy beültetett eszközei vannak, amelyek folyamatosan figyelik az életjeleket és az immunjelzőket; Az egészségügyi adatok egyben oltási útlevélként is szolgálnak, és elengedhetetlenek a közéletben való részvételhez. A robotika az emberi kontaktus hiányait is pótolja – az autonóm szállító drónoktól a háztartási segítő robotokig, amelyek vásárlást, takarítást vagy akár alapvető orvosi feladatokat is elvégeznek, hogy minimalizálják az ember-ember közötti kitettséget. A történetben a brazil főhős orvos, aki mesterséges intelligencia eszközökkel távolról kezeli a betegeket, míg a kínai főszereplő lelkes játékos. Egy online többjátékos játékban "találkoznak", amely kevert valóságot és valós adatokat ötvöz. Virtuális interakcióik révén láthatjuk az elszigeteltség érzelmi terhét, de ugyanakkor az újszerű intimitási formákat is, amelyeket a technológia kínál – például a virtuális randevúik fantasztikus tájakban szinte valóságosnak tűnnek a haptikus visszacsatolásoknak köszönhetően. A "Kontaktus nélküli szerelem" cím azt tükrözi, hogyan talál utat a romantika a fizikai elkülönülés ellenére is, mesterséges intelligencia által közvetített élményeken keresztül. Azonban egy mellékszál egy új COVID-variáns járványt

érint, ami jól mutatja, milyen gyorsan reagál az MI-vezérelt biotechnológia: a laborok mesterséges intelligencia gyógyszerfeltárást (például az AlphaFoldot) használnak a vírus modellezésére és néhány napon belül testreszabott vakcinát előállítására. A társadalom 2041-ben lényegében megtanult egy endemikus vírussal együtt élni az automatizálásra, a gyors mesterséges intelligencia epidemiológiára és a szigorú, adatvezérelt egészségügyi rendszerekre támaszkodva.

Lee esszéje ebben a történetben az MI-re fókuszál az egészségügyben és az orvostudományban, különösen a járdémia által felgyorsítva. Arról beszél, hogyan forradalmasítják az olyan MI rendszerek, mint a *DeepMind AlphaFoldja* (amely megoldotta a fehérjehajtogatást), forradalmasítják a gyógyszerfejlesztést – 2041-re az MI rutinszerűen tervezhet vírusellenes szereket vagy vakcinákat bármilyen új kórokozóra válaszul. A robotika és az MI-vezérelt automatizálás "előrelépést" mutatott be a járványok idején: a könyv elképzelte, hogy mindenütt jelenlévő szolgáltató robotok szállítják, fertőtlenítik a helyiségeket, vagy akár idősgondozást biztosítanak, így csökkentve a kockázatos emberi kontaktusokat. Lee kiemeli, hogy a COVID-19 katalizátorként működött a biotechnológiai innováció és a teleprezentáció számára. Ebben a jövőben az MI diagnosztikával működő távmedicina gyakori – a betegek otthon is MI-vezérelt orvosi vizsgálatot végezhetnek, ahol az életjeleket viselhető eszközökön keresztül továbbítják, és algoritmusok szűrik az anomáliákat. Az elemzés *etikai dilemmákat is érint*: egy 24/7 egészségügyi megfigyeléssel rendelkező világ életet menthet, de a személyes magánélet rovására. Ha folyamatosan nyomon követik a pulzusodat, a vérkémiai összetételedet és a helyedet, akkor kié az adatok és vissza lehet használni? A történet forgatókönyve utal a *közegészségügy és az egyéni szabadság közötti kompromisszumra*. Lee olyan fogalmakat vezet be, mint a szövetséges tanulás és a biztonságos enklávék, mint módok, amelyek az MI felhasználhatja a személyes egészségügyi adatokat a magánélet megőrzése mellett. A kapcsolatok oldalán ez a fejezet kérdéseket vet fel a virtuális valóságról és az emberi kapcsolatról: vajon a virtuális érintés és az AI avatárok valóban helyettesíthetik a személyes kapcsolatot? Milyen pszichológiai hatások merülnek fel, ha a társas életünket a technológia közvetíti (jó vagy rossz értelemben)? Összességében a "Kontakt nélküli szeretet" egy reális képet fest egy közeli jövőről, amelyet egy tartós járvánnyal formál – az MI-vel gyorsított egészségügy, a robot segítők és a gazdag virtuális világok együtt tartják fenn a társadalmat és a kapcsolatokat életben tartják, bár átalakított módon.

(Elemzési témák: MI az egészségügyben és biotechnológiában, AlphaFold által vezérelt gyógyszer-/vakcina felfedezése, orvosi robotika, járdémiás automatizálás, viselhető eszközök és egészségügyi adatok, teleprezentációs technológia.)

5. "Kísérteties bálványom" – Immersive Mixed Reality és Virtuális Sztárok (Tokió, 2041)

Ez a történet 2041-ben merül be az XR (kiterjesztett valóság) szórakoztatás világába – ahol a valóság és a virtuális közötti határok szinte teljesen eltűntek. Tokióban játszódik, és egy elkötelezett zenerajongót követ nyomon, aki beleszeret egy virtuális popsztárba – egy mesterséges intelligencia által generált virtuális idolba, aki csak vegyes valóságos előadásokban "létezik". Az általános AR szemüvegeknek és neurális felületeknek köszönhetően 2041-ben az emberek szabadon bővíthetik a gazdag virtuális tartalmat a fizikai környezetükre. A főszereplő élő koncerteken vesz részt, ahol a popidol tökéletesen realiztikus hologramként jelenik meg a színpadon, olyan karizmával énekel és táncolva, amely bármely emberi előadóval vetekedhet. Az AR-ben egyénileg is találkozik "vele" – például az idol avatár elkísérheti sétára, az AR lencsein keresztül jelenik meg mellette, és reagál a szavaira (fejlett természetes nyelvi

mesterséges intelligencia által működtetve). Idővel a rajongók megszállottsága mélyül. A virtuális idol mindig elérhető, tökéletesen kialakítva, hogy a rajongókat kielégítse (személyisége szó szerint a maximalizált elköteleződésre van programozva), és még *kísértetiesen* élethűen is élethű. A fiatal ember visszahúzódik a valódi barátoktól és felelőségektől, ahogy több időt tölt az idol magával ragadó XR világában. A történet pszichológiai válságba vezet: mivel nem tudja megkülönböztetni a virtuális kényelmet a valóságtól, kockáztatja, hogy elveszíti önmagát. *"Olyan, mintha nyitott szemmel álmodnál"* – idézi a könyv – egy csábító menekülés, ami addiktívvá válik. A "My Haunting Idol" cím arra utal, hogy ez az AI híresség szó szerint kísérti a rajongók elméjét, szemlélteti az ultra-realisztikus virtuális élmények lehetséges pszichológiai terhét.

A kapcsolódó elemzésben Lee áttekinti azokat a technológiákat, amelyek lehetővé teszik ezt a helyzetet: virtuális valóság (VR), kiterjesztett valóság (AR), vegyes valóság (MR) – összesen kiterjesztett valóság – valamint az agy-számítógép interfészeket (BCI). 2041-re jóslása szerint az XR *"mindenütt"* és zökkenőmentes lesz. A mai esetlen VR headsetek és primitív hologramok könnyű szemüvegekké vagy akár idegi implantátumokká fejlődnek, amelyek nagy pontosságú 3D karaktereket és jeleneteket képesek kivetíteni a látómezőnkbe, megkülönböztethetetlenül a fizikai tárgyaktól. Ezekkel a környezetekkel való interakció több érzéket is bevon: nemcsak látást és hallást, hanem érintést (haptikus ruhákon keresztül), és esetleg közvetlen idegi stimulációt (fejlett BCI) a textúra vagy hőmérséklet közvetítésére. Lee párhuzamokat vont a meglévő technológiával: például a Microsoft HoloLens már vegyes valóságot kínál, és bár olyan termékek, mint a Google Glass a 2010-es években kudarcot vallottak, az alapvető technológia folyamatosan fejlődött. A 2040-es évekre elképzelhető egy egész "virtuális idol" iparág – valóban, még 2023-ban is vannak korai virtuális influencerek és popsztárok. Lee a társadalmi következményekről beszél: a szórakoztatás hihetetlenül személyre szabott és magával ragadó lesz (az emberek igény szerint fantáziavilágokat is létrehozhatnak), de fennáll a menekülés, függőség és az identitás elmosódásának kockázata. Ki birtokolja az AI idol dalainak vagy személyiségének jogait? Mi van, ha az emberek inkább a virtuális társakat részesítik előnyben a valódiak helyett – ez aláás-e a társadalmi összetartást? A megszállott rajongó története egy figyelmeztető történet a *valóság elvesztéséről*. Lee hangsúlyozza az etikus tervezés szükségességét az ilyen rendszerekben – például az MI idoloznak kellene emlékeztetniük a felhasználókat arra, hogy "nem vagyok valódi", vagy korlátozzák az interakciókat, hogy megakadályozzák az egészségtelen kötődést. Az elemzés kiemeli a pozitív oldalt is: az XR lehetővé teszi a kreatív kifejezést és a művészethez való globális hozzáférést, ahogy korábban soha nem volt (bárki, bárhol virtuális módon is megtapasztalhatja a tokiói koncertet). Végső soron a "My Haunting Idol" arra kér minket, hogy gondoljuk meg, hogyan lehet egyensúlyba hozni az elmélyedő MI szórakoztatás csodáit a mentális egészséggel és a valósághoz való megalapozással.

(Elemzési témák: virtual/kiterjesztett/vegyes valóság, agy-számítógép interfészek, virtuális hírességek, valamint az elmélyedő MI etikai/társadalmi kérdései.)

6. "A szent sofőr" – Autonóm járművek és ember–mesterséges intelligencia együttműködés (Colombo, 2041)

Ebben a fejezetben az önvezető autókra és a teljesen autonóm közlekedésre való összetett átmenetre összpontosítunk. A történet Colombóban, Srí Lanka-ban játszódik, ahol a kaotikus és kiszámíthatatlan városi forgalom háttérét nyújtja. 2041-re sok jármű az úton autonóm lesz (mesterséges intelligencia "vezetőkkel" felszerelve), de nem mindegyik. A főszereplő egy veterán emberi sofőr – egykor híres helyi buszsofőr volt, aki ügyességéről és merészségéről

ismert –, aki most egy szokatlan munkában dolgozik: ő egy távoli "távvezető", aki segít autonóm járműveknek nehéz helyzetekben. Lényegében, amikor egy MI vezetési rendszer olyan szélső esettel találkozik, amit nem tud kezelni (például hirtelen földcsuszamlás vagy bonyolult építkezési kitérő), egy emberi tele-vezetőt jelöl, hogy virtuálisan vegye át az irányítást. A történet egy konkrét eseményt dramatizál: egy teljesen autonóm autó veszélyes árvíz áldozata lesz egy vidéki úton, és az MI nem tudja megmenteni, hogyan mentheti meg az utasokat. Az emberi sofőrünk Colombóban teleoperatív szerkezeten keresztül "beugrik" (hasonlóan ahhoz, ahogy egy országban a drónpilóták távolról irányítanak drónokat), és tapasztalt ösztöneivel manuálisan navigálja az autót biztonságba. Ez a hősiesség – lényegében egy ember, aki a világ másik feléről menti meg embereket egy hálózatos járművel – megkapja a "szent sofőr" becenevet. Ez azt is hangsúlyozza, hogy még 2041-ben sem váltotta ki teljesen az emberi ítélőképességet minden helyzetben. A társadalom határfázisban van: az 5. szintű teljes autonómia közel van, de nem univerzális. A történet olyan feszültségeket vizsgál, mint a közbizalom (egyesek nem hajlandók autonóm autóval utazni balesetek után), valamint a munkaerő-problémákat (a legtöbb emberi sofőr elmozdul, kivéve azokat, mint a főszereplő, akik távirányítóként képezik át).

Lee elemzése az "A szent sofőr" című könyvben az autonóm járművek (AV) állapotát és az 100%-os önvezető biztosítás elérésének akadályait tárgyalja. Ő magyarázza a SAE automatizálási szinteket (L0-tól L5-ig) – ahol az L5 azt jelenti, hogy egy olyan jármű, amely minden körülmények között képes önmagát vezetni, kormánykerék nélkül. A 2020-as évek elejére kontrollált beállításokban elértük a L2 vagy L3 szintet; A könyv szerint 2041-re sok út és okosváros elfogadja az L4+ autonómiát. Lee azonban rámutat kihívásokra: a kiszámíthatatlan környezetek, gyenge infrastruktúra vagy ritka "fekete hattyú" események (például a villámáradás) még mindig megzavarhatják az MI-t. A történet emberi tele-driver-forogatókönyve egy lehetséges megoldás – egyfajta *emberi biztonsági mentés* a szélsőséges esetek kezelésére. Ez tükrözi a drónháború lefolytatását (Nevadában pilóták irányítanak drónokat konfliktuszónákban), és egy olyan jövőt sugall, ahol egyes sofőrök irányítóközpontokban ülnek, és felügyelnek robotaxisok flottáját. Az esszé az AV-t lehetővé tévő technológiákat vizsgálja: a fejlesztett érzékelők (LIDAR, kamerák), V2X kommunikáció (járművek egymással és városi infrastruktúrával beszélnek), valamint jobb döntéshozatali algoritmusok. 2041-re a járművek valószínűleg okosvárosi rendszereken keresztül koordinálódnak a forgalom optimalizálása érdekében. Etikai kérdésekről is szó esett: a klasszikus "villamosprobléma" – hogyan mérlegelje az AI-nak az utasok életét a gyalogosok életével egy elkerülhetetlen baleset esetén? A történet ezt érinti, amikor leírja, miért győzi meg néha az emberi intuíció a merev MI logikát válságokban. Lee megjegyzi, hogy a társadalomnak döntést kell hoznia az AV döntéshozatal szabályozásáról és erkölcsi kereteiről. Ott van a társadalmi-gazdasági szempont is: milliók hivatásos sofőr (kamionosok, taxisok) veszíthetnek el állásukat; Újraképzési programok (például a sofőrök teleoperátorra vagy karbantartó technikussá alakítása) elengedhetetlenek. Bátorító módon 2041-re az összehatás biztonságosabb utak és hatékonyabb közlekedés lehet, mivel az MI gyorsabban reagál, mint az emberek, és csökkenti a baleseteket (egy szereplő csodálkozik, hogy a közlekedési halálesetek száma drasztikusan csökkent). A "The Holy Driver" végső soron egy partnerségi modellt mutat be – az emberek és az MI mindketten azt teszik, amiben a legjobban tudnak –, utalva arra, hogy még ha az AI növekszik is, az emberi szakértelemnek továbbra is szerepe lesz kritikus helyzetekben.

(Elemzési témák: autonóm járművek technológia, L0–L5 autonómia szintek, integráció az okosvárosokkal, valamint az önvezető autók etikai és társadalmi szempontjai.)

7. "Kvantumnépirtás" – MI fegyverek és a technológia sötét oldala (München & global, 2041)

Az egyik legbaljósabb látomás, a *"Kvantumirtás"* egy thriller, amely az MI-vezérelt hadviselés és a kvantumszámítás rossz kezekben való összefonódását képzei el. A történet egy örült európai technológus körül forog, aki tömegpusztítási tervet dolgoz ki – lényegében népirtás kísérletét – autonóm fegyverekkel. A kvantumszámítás áttörései segítségével sikerül feltörnien az összes titkosítási és biztonsági rendszert, amelyek különböző katonai és infrastrukturális hálózatokat védenek. Ez azt jelenti, hogy képes átvenni az AI-val működő harci drónok és rakétarendszerek flottáit. 2041-ben sok ország autonóm drónokat használ védelemre; Ez a gonosztevő átprogramozza őket, hogy egy adott populációt célozzanak meg. Közben egy kazahsztáni hacker válik a váratlan hőssé, aki az idővel versenyez, hogy megakadályozza a katasztrófát azzal, hogy ellenhackelte az AI "vírust", amelyet az örült tudós szabadított ki. A narratíva nagy tétet foglal magában: fegyverrel elkészített drónrajok sötétítik az eget, globális pánik a kommunikáció meghibásodásával (például a kvantumhackelés megtöri a blokklánc-alapú pénzügyeket, ami a Bitcoin és a digitális infrastruktúra összeomlásához vezet). A "kvantumnépirtás" kifejezés példátlan méretre utal – milliókat ölhetnek meg nem egy nukleáris bomba, hanem a hálózatba kapcsolt MI-fegyverek által a civilek ellen, amelyeket a megtörhetetlen kvantumkód-feltörés tesz lehetővé. Végül a hackerhős talál egy kiskapukat, amellyel leállítja a szökött MI-t, elkerülve a végítéletet, de az eset éles figyelmeztetésként szolgál.

Lee esszéje két fő témába merül: a kvantumszámítás és az autonóm fegyverek. A kvantumszámítógépek, ellentétben a klasszikusokkal, a kvantumfizikát használják, hogy bizonyos számításokat csillagászatosan gyorsabban végezzenek. 2041-re elképzelhető, hogy egy erős kvantumszámítógép szinte azonnal megtörheti a mai szabványos kriptográfiát (például az RSA vagy a Bitcoin SHA-256-ját). Lee ezt a fenyegetést magyarázza, és hangsúlyozza, hogy kvantumellenes titkosítás szükséges, mielőtt ezek a gépek belépnének. Az MI fegyverek terén arról beszél, hogyan használják az MI-t katonai kontextusokban – az intelligens megfigyeléstől az autonóm drónokig és őrségi fegyverekig. A történet forgatókönyve egy egzisztenciális kockázatot emel alá: ha az MI által vezérelt fegyverek széles körben és hálózatba kapcsolódnak, egyetlen sértés vagy rosszindulatú MI ellenünk fordíthatja őket. Ez összhangban van a valódi vitákkal, amelyek a "gyilkos robotok" betiltásáról és az emberi felügyelet biztosításáról a halálos döntésekben. Lee megjegyzi, hogy 2041-re az autonóm fegyverek valóban olcsók és bőségesek lehetnek (például mesterséges intelligencia mini-drónrajok, amelyeket egy terrorista telepíthet). A kvantumhackeléssel való kombináció különösen erős – egyetlen rendszer sem lenne biztonságos. Az elemzés kiemeli, hogy *maga az MI* is fegyver lehet (például MI által vezérelt kibertámadások, háborús deepfake propaganda), és hogy globális kormányzás szükséges az MI fegyverkezési verseny megelőzéséhez. A fejezet valószínűleg történelmi analógiákat említ: ahogyan a nukleáris proliferációhoz szerződésekre volt szükség, úgy az MI és a kvantumtechnológia nemzetközi megállapodásokat igényel a katasztrófális visszaélések elkerülése érdekében. Lee hangneme itt óvatos, hangsúlyozva ezt a víziót, mint egy "figyelmeztető történetet" arról, hogyan lehet a technológia exponenciális ereje a pusztulás felé fordítani. Az autonóm fegyvereket *"egzisztenciális fenyegetésnek nevezi az egész világra a rossz szereplők kezében."* Ugyanakkor a könyv azt is sugallja, hogy az ilyen helyzetek tudatossága az első lépés azok megelőzésében. A hős sikere arra utal, hogy megfelelő biztonsági intézkedésekkel (pl. MI-k keményen kódolt etikája, biztonsági leállások, kvantum utáni titkosítás) az emberiség képes megtörni ezeket a rémálmokat. Összefoglalva, a "Kvantumirtás" az MI végső etikai tétjét tárgyalja: a túlélést. Azt állítja, hogy proaktívan

biztosítani kell, hogy az MI és a kvantumfejlesztések globális etika vezérelje, és ne legyenek könnyen eltéríthetők vagy helytelenül telepíthetők.

(Elemzési témák: kvantumszámítás (és annak hatása a titkosításra/biztonságra), blokklánc sebezhetőség, autonóm drónok és fegyverek, mesterséges intelligencia által vezérelt hadviselés, egzisztenciális kockázatok.)

8. "A munkahelymegmentő" – Munka az MI korában és a 3Rs megoldás (Szilícium-völgy, 2041)

Ez a fejezet az MI automatizálás okozta gazdasági felfordulást tárgyalja, de egy optimista csavarral. A történet az Egyesült Államok technológiai központjában (Szilícium-völgyben) játszódik, és egy fiatal nőt követ, aki "állásátosztó ügynökként" dolgozik. 2041-re az MI és a robotok rengeteg hagyományos munkahelyet szüntettek meg – a teherautó-sofőröktől a könyvelőkig – nagy elmozdult munkaerőt teremtve. Erre válaszul egy új iparág jelent meg: olyan ügynökségek, amelyek segítenek átképezni és az elbocsátott embereket új szerepkbe helyezni, ahol még szükség van emberi készségekre. A főszereplő lényegében az MI korszak karriertanácsadója, aki az ügyfeleket vezeti a Kai-Fu Lee által "3R-nek" nevezett módon: Újratanulás, Újrakalibrálás, Reneszánsz. Például segíthet egy volt gyári munkásnak *újratanulásban* egy új napelem-telepítési munkához (egy olyan területen, ahol még kereslet van az emberi munkaerőre), *átalakítani* elvárásait és gondolkodásmódját az MI mellett való munkához (például technikussá válni, aki egy MI rendszert felügyel, nem pedig kézi munkát végez), és megtalálja a *reneszánszát* – egy új karrier, amely egyedi emberi tehetségeket, mint a kreativitást vagy az emberi empátiát érinti. A történet valódi kihívásokat jelent: néhány munkanélküli ügyfél depressziós vagy ellenálló a változásnak; Nem elég új munka mindenkinek; Politikai feszültségek fortyognak az univerzális alapjövedelem (UBI) rendszerek miatt, amelyek támogatják azokat, akik nem találnak munkát. A főszereplő etikai dilemmával szembesül, amikor rájön, hogy ügynöksége, a befektetők nyomása alatt, csak a leg"áttelepíthetőbb" jelölteket választja ki a sikermutatók fenntartására, miközben csendben feladja másokat. El kell döntenie, hogy lefújja-e a sípát, vagy alkalmazkodik. Útja során a történet kiemeli a veszteséget és a reményt is: milliók veszítették el karrierjüket, mégis újfajta munkák jelentek meg (néhányik elég kielégítő), és a társadalom olyan biztonsági hálókkal kísérletez, mint az UBI.

Elemzésben Lee az MI által vezérelt munkahelyek elmozdulására és annak kezelésére összpontosít. Megjegyzi, hogy *"az MI kifejezett célja, hogy átvegye az emberi feladatokat, ezzel megsemmisítve a munkahelyeket"* – ellentétben a korábbi technológiákkal, amelyek gyakran annyi munkahelyet teremtettek, mint amennyit tönkretettek. 2041-re ez a tendencia teljes erővel beindul: bármilyen *rutinszerű, ismétlődő* munka valószínűleg automatizált lesz. Ez nemcsak gyári vagy irodai munkát foglal magában, hanem néhány fehérgalléros munkát is magában foglal (mesterséges intelligencia alapvető jelentések írása, betegségek diagnosztizálása stb.). A könyv azt állítja, hogy *"újrágondoljuk a munka fogalmát"*, és be kell vezetnünk olyan intézkedéseket, mint az újraképzési programok, az UBI és az emberközpontú készségek oktatása. Lee "3R" keretrendszere – amelyet a történet is bemutat – az egyének és társadalmak számára egy stratégiaként jelenik meg:

- Újratanulás – folyamatosan frissítsd a képességeidet az AI fényében (például megtanulsz AI eszközökkel dolgozni, vagy áttérsz nehezebben automatizálható területekre).

- Újrakalibrálni – igazítsuk a politikákat és a személyes elvárásokat, felismerve, hogy a karrierpályák és identitások gyakrabban változnak az AI miatt (például az élethosszig tartó tanulás elfogadása, és a döntéshozók ösztönzők létrehozása új munkaerőpiacok számára).
- Reneszánsz – különös hangsúly az egyedi emberi tulajdonságokra, mint a kreativitás, összetett problémamegoldás, személyes empátia és művészi képesség – olyan területekre, ahol az MI küzd, és amelyek gazdagítják az emberi életet. Lee szerint egy MI-tel telített gazdaságban olyan munkahelyek, amelyek emberi kapcsolatot (tanárok, ápolók, vállalkozók, művészek) vagy új kreatív törekvések, virágozhatnak, ami egyfajta "emberi reneszánszt" jelenthet.

Az elemzés az UBI-t és az új gazdasági modelleket is mérlegeli. Azok számára, akik teljesen kiűzöttek, az UBI biztonsági hálót biztosít – a könyv szerint 2041-re néhány ország vagy állam sikeresen kísérletezett az alapjövedelemmel, amelyet talán az MI termelékenységére kiadott adókból finanszíroznak. Lee elismeri, hogy az MI drasztikusan növelheti az egyenlőtlenséget (a technológiai óriások és az AI tulajdonosok élvezik a legtöbb előnyt), ezért a társadalmi beavatkozás kulcsfontosságú a tartós alsóosztály megelőzéséhez. Érdekes módon a történet és az elemzés hangsúlyozza, hogy a munka sok embernek ad értelmet, így még ha az UBI anyagi szükségleteket is fed le, az emberek értelmet keresnek – ezért fontos, hogy segítsenek az embereknek új hivatásokat találni (a főszereplő szerepe, mint "munkamentő" szimbolikusan megmenti az emberek céltudatát, nem csak a jövedelmüket). Összefoglalva, a "The Job Savior" egy képet fest a munkaerőpiacról 2041-ben, amely viharos, de nem reménytelen. Proaktív alkalmazkodással – képzéssel, szociális támogatással és az emberi erősségekre helyezett hangsúllyal – a munkaerő átalakulhat, sőt, még kielégítőbb utakat is találhat, ahelyett, hogy tömeges munkanélküliségnek esne át.

(Elemzési témák: AI munkahelyautomatizálás és kiszorítás, olyan stratégiák, mint az UBI és az újraképzés, mit nem tud az MI (csak emberi készségek), valamint a "3R" keretrendszer a társadalmi alkalmazkodáshoz.)

9. "A Boldogság Szigete" – MI, adatok és az emberi boldogság keresése (Arab-tenger, 2041)

Ez a gondolatébresztő történet egy mesterséges boldogság kísérletére fókuszál. Egy gazdag közel-keleti uralkodó (feltételezhetően a Perzsa-öböl régiójában) egy eldugott luxusszigetközösséget hoz létre, amely az emberi boldogság maximalizálása érdekében az MI segítségével. Ezen a "Boldogság Szigetén" a lakosok (kezdetben milliárdos önkéntesek) beleegyeznek, hogy *minden* személyes adatukat – érzelmi, fizikai, társas – átadják a sziget MI-rendszereinek, amelyek aztán az élet minden aspektusát úgy alakítják ki, hogy örömeiket szolgálja. Az MI mindent figyel: biometrikus szenzorok valós időben mérik az egyes személyek endorfin-, szerotonin- és dopaminszintjét; az okos otthonok rögzítik tevékenységeiket és hangulati válaszaikat; A társas interakciókat elemzik. Ezen adatok felhasználásával az MI személyre szabott élményeket szervez – a szórakozástól és ételtől a társas találkozásokig –, mindezek célja, hogy minden egyes egyén boldogságmutatói optimálisak legyenek. Például, ha egy lakó magányosnak érzi magát, a rendszer finoman bevezethet egy kompatibilis barátot az időbeosztásába; Ha az izgalom csökken, az meglepetés izgalmat okozhat (elméletileg etikusan). Kezdetben ez az adatvezérelt utópia magas boldogsági mutatókat hoz létre – a résztvevők örömmel töltik el, milyen tökéletesen hangolva tűnik az élet az ő preferenciáikhoz. Azonban ahogy telik az idő, repedések jelennek meg. A történet egy lakót követ, aki a folyamatos kényeztetés és kényelem ellenére kezd ürt érezni. Rájön, hogy hiányzik a *cél* és a

személyes fejlődés – az MI annyira a küzdelem megszüntetésére koncentrál, hogy az élet hátborzongatóan nyugodttá és értelmetlenné vált. Néhány lakó unalmassá vagy hedonizmusba (még szerhasználatba is), mert amikor minden igényt könnyedén kielégítenek, elveszítik a motivációt. A sziget nagy kísérlete tehát egy paradoxonnal néz szembe: a boldogság, mint mérhető cél, nem feltétlenül jelent egy beteljesülést. A csúcsponton a monarcha összeáll az AI alkotóival, hogy finomítsák az algoritmust – esetleg kihívásokat vezetnek be vagy önzetlen viselkedést ösztönöznek –, kiemelve, hogy *az igazi boldogság több, mint pusztán a maximált örömmérők.*

Lee elemzése az MI-ről és a boldogság tudományáról, valamint az adatvédelmi és az ilyen vállalkozáshoz szükséges adatokkal kapcsolatos kényes kérdésekről szól. Megjegyzi, hogy a boldogság szubjektív és sokrétű – amit még a legokosabb MI is nehezen tudna pontosan meghatározni vagy mérni. A történet aláhúzza, hogy egy MI képes minden adatunkat (IoT eszközök, közösségi média, egészségügyi adatok – *"mindent"*, ahogy a könyv mondja), és mégsem értheti meg a boldogság jelentését, mert részben filozófia. Lee az ősi bölcsességre utal: sok bölcs mondta, hogy a boldogság belülről jön, nem külső ingerekből – ezt a felismerést az MI-sziget kemény úton tanulja meg. A technikai oldalon ez a fejezet azt vizsgálja, hogyan lehetséges 2041-re példátlan módon összegyűjteni a személyes adatokat. Azonban ez etikai és jogi akadályokba ütközik, például adatvédelmi törvények (például Európa GDPR-je tiltja az ilyen átfogó személyes adatgyűjtést). Az elemzés olyan fogalmakat vezet be, mint a szövetséges tanulás és a homomorf titkosítás, amelyek lehetővé teszik az MI számára, hogy személyes adatokból nyers adatokat feltárva nyers adatot nyerjen – ami potenciális kompromisszum a magánélet megőrzése érdekében, miközben az adatokat jóra használják. Tárgyalnak továbbá a személyes MI-asszisztensek ötlete is, amelyek a mi érdekeinknek szolgálhatnak – például egy olyan MI, amely tanácsot ad az életválasztásokban, amelyek összhangban vannak a boldogságoddal, de csak akkor, ha megbízik benne az intim adatokkal. Fontos, hogy Lee elmélyül, *mit kellene az AI-nak optimalizálnia?* Ha azt mondod az MI-nek, hogy maximalizálja a boldogság kémiai proxyját, akkor előfordulhat, hogy wireheadinget kapsz (MI dopaminnal pumpálja az agyadat). Ehelyett talán az MI-nek segítenie kellene az embereknek a cél, kapcsolatok és erény követésében – kevésbé kézzelfogható, de végső soron kielégítőbb célokat. A sziget kudarca azt tanítja, hogy egy rosszul választott mérőszám (még a "boldogság" is) vak optimalizálása visszaüthet – ez a lecke az MI összehangolódásáról: rendkívül óvatossá kell lennünk az MI-nek adott célok meghatározásában, különben nem kapjuk meg, amit *kérünk*, nem pedig azt, amit igazán akarunk. Összefoglalva, a "Isle of Happiness" megkérdőjelezi azt a nézetet, hogy az MI meg tudja vagy kellene megoldania az emberi állapotot, és hangsúlyozza az emberi cselekvési képesség és a rejtély megőrzését a jólét keresésében.

(Elemzési témák: MI és emberi boldogság, adatvédelem (GDPR) és személyes adatok aggregálása, adatvédelmet megőrző számítás (szövetségi tanulás, TEE), valamint a szubjektív célok optimalizálásának filozófiája.)

10. "Álmodni a bőségeről" – A bőség kora és az új gazdasági modellek (Ausztrália, 2041)

Az utolsó fejezet egy olyan világot ábrázol, ahol az MI és más technológiák elindították a szűkös "bősége" korszakát. A történet Ausztráliában, körülbelül 2041-ben játszódik, és azt képzelet el, hogy az MI, a megújuló energia, az automatizálás és az anyagtudomány áttörései drámaian csökkentették az áruk és szolgáltatások árát. A tiszta energia (nap, szél, fúzió stb.) annyira bőséges, hogy az áram szinte ingyenes. Az MI-automatizált gyárak és 3D nyomtatók alapvető termékeket – élelmiszert, ruházatot, építőanyagokat – állítanak elő szinte nulla

határköltiséggel. A legtöbb állampolgár "*Basic Living Support*" (egy univerzális *Basic Life Card révén*) kap, amely minden alapvető szükségletet és kényelmes életmódot fedez. A szegénységet és az éhséget gyakorlatilag megszüntették ez a techno-gazdasági bőség által. Felszínesen ez úgy néz ki, mintha az utópia megvalósulna: mindenkinek van szabadideje, senki sem fél az éhezéstől vagy hajléktalanságtól, és a technológia mindenkinek bőven szolgál. A történet egy fiatal őslakos ausztrál nőt, Keirat követ, aki ebben a társadalomban él. Ahelyett, hogy a túlélésért dolgozna, Keira közösségi szolgálatban (idősek gondozója) és kreatív tevékenységekben dolgozik, társadalmi elismeréseket szerezve. A kormány (vagy az MI adminisztrátorok) bevezettek egy "Jukurrpa" nevű rendszert – amely az őslakos álomidő koncepciójáról nevezett el –, amely a polgárokat a társadalmi viselkedésért jutalmazza egy "Moola" nevű hírnév valutával. Mások segítése vagy önkénteskedése Moola pontokat eredményez, amelyeket AI karkötővel követnek, hogy ösztönözzék a hozzájárulást egy olyan világban, ahol a hagyományos munka opcionális. Keira értelmet talál abban, hogy másokról törődjön, Moolát nem vagyoni miatt gyűjt (az alapvető dolgok ingyenesek), hanem jóindulatú státuszként. Azonban a történet feltárja, hogy még bőséges állapotban is fennállnak az emberi kihívások. Néhányan elkezdik játszani a Moola rendszerrel (játészó "jócselekedeteket" végeznek pontok szerzésére), ami státuszversenyzővé változtatja. Mások, különösen a fiatalok, céltalansággal küzdenek – gazdasági nyomás nélkül néhányan unatba vagy nihilizmusba, sőt szerhasználatba esnek (ez a helyzet visszatükrözi az aggályokat, hogy "a grind" nélkül az emberek elveszíthetik a célt). A társaság olyan megoldásokat kísérletez, amelyek motiváltak és boldogok maradnak az embereknek – például maga a Jukurrpa rendszer. Keira útja magában foglalhatja a családok leleplezését vagy a rendszer igazságosabbá finomítását, ami azt mutatja, hogy még egy majdnem utópia is etikus fenntartást és folyamatos fejlesztést igényel. A fejezet címe, "Álmodni a bőségeről" arra utal, hogy a bőség régóta tartott álom volt, de annak eléréséhez még mindig bölcsesség kell a következmények kezeléséhez.

Záró elemzésében Lee visszatekint azokon az új gazdasági modelleken, amelyeket a szűkös világ megkövetelne. A hagyományos kapitalizmus és a bérért dolgozó logika meghibásolható, amikor az MI és az automatizálás minimális emberi munkaerővel a legtöbb értéket hozza létre. Azt javasolja, hogy 2041-re talán megjelennek a "*világ magjai, amely a szűkösségből a bőségbe mozog*", ami újjá kell találni az erőforrások elosztását és ösztönző társadalmi részvételt. Az olyan ötletek, mint az univerzális alapjövedelem (vagy a történetben az Alapvető Életkártya) kulcsfontosságúvá válnak – biztosítva, hogy mindenki profitáljon az AI termelékenység-növekedéséből. Lee megjegyzi, hogy ha az MI és a robotika szinte szabadon tudnak vagyont termelni, "*a jelenlegi közgazdasági elmélet már nem fog érvényesülni*" – olyan mutatók, mint a GDP vagy a munkaidő, talán nem számítanak. Ehelyett a fejlődést az életminőség, az oktatás, a kreativitás és a társadalmi jólét javulásával mérjük. A történet Moola rendszere példája annak, hogy a pénz nem motiválja a pozitív cselekvéseket. Lee a lehetséges buktatókat tárgyalja: még a jó szándékú rendszerek is tökéletlenek lesznek és kizsákmányolhatóak (ahogy azt az emberek Jukurrpa játékában is láthatják). A kulcstéma, hogy *a technológia nem oldja meg automatikusan a társadalmi problémákat* – az emberi értékeknek és a kormányzásnak kell irányítaniuk, hogyan használjuk ezt a bőséget. Lee itt érinti a szingularitás fogalmát: vajon ez a bőség vezet elő a technológiai szingularitást? Személy szerint kétféle, hogy 2041-re valódi Szingularitás (szökött szuperintelligencia) fog bekövetkezni, de elismeri, hogy a világ drámaian más lesz. Az elemzés valószínűleg reménytelen, mégis óvatos jegyben ér véget: *ha bölcsen kezeljük az AI előnyeit*, megszüntethetjük a szélsőséges nélkülözést és nagy léptékben felszabadíthatjuk az emberi potenciált (több ember szabadon tanulhat, művészetet alkothat, gondoskodhat a közösségről). Azonban, *ha nem tervezünk előre*, még az utópikus technológia is disztópikus kimenetekhez vezethet (önelégültség, jelentésvesztés, új státuszbeli vagy befolyási egyenlőtlenségek). Így a "*Dreaming of Plenitude*"

a könyv üzeneteit összekötő csúcspontként szolgál: az MI lehetővé teszi a virágzó jövőt, de az emberi cselekvés, az etika és a társadalmi rendszerek innovációja határozza meg, hogy valóban fényes-e ez a jövő.

(Elemzési témák: bőséges/poszt-szűkös gazdaság, a munkára és a motivációra gyakorolt hatások, a pénz és ösztönzők újragondolása, alapjövedelem, a pénz jövője, valamint a szingularitás vitára egy áttekintés.)

Összekapcsolódó szálak és végső gondolatok

A tíz történet után Kai-Fu Lee egy összefoglalót ad, kiemelve az átfogó témákat. Megjegyzi, hogy az *AI 2041* forgatókönyvei nem túlrekedett sci-fi, hanem "felelős és valószínű forgatókönyv-halmaz" – sőt, úgy becsülte, hogy minden technológiai fejlesztésnek legalább 80%-os eséllyel valósul meg 2041-re. A könyv utolsó üzenete arra szól, hogy ébredjenek fel az MI potenciáljára és kockázataira, és proaktívan alakítsák azt. Érdekes módon Chen Qiufan záró története (amely a 10. fejezet részeként van beágyazva) ügyesen köti össze a korábbi történetek szereplőit vagy eseményeit, felfedve a narrátor meglepetését, amely összeköti az összes történetet. (Nem áruljuk el a csavart, de kielégítő égségérzetet ad a gyűjteménynek.)

Több kulcsfontosságú téma rezonál a víziókon:

- Emberi cselekvés és etikai összehangolás: Szinte minden történet szembeállítja az MI hatalmas képességeit azzal, hogy az emberi értékek vezérelné azt – legyen szó a biztosítási előítéletek visszaszorításáról (1. történet), az igazság biztosításáról a médiában (2. történet), vagy arról, hogy mit *érdemes* optimalizálni (9. történet). A szerzők ismételten hangsúlyozzák, hogy maga a technológia *nem alapvetően erkölcsös*; az eredmények attól függenek, hogyan tervezzük és használjuk azt. Lee gyakran semleges eszközként írja le az MI-t – *"egy objektív technológiát, amely csak az emberek általi használat révén szerez etikai értéket"* –, figyelmeztetve, hogy ha az MI kárt okoz, az az emberi félreállítás vagy visszaélés miatt történik. A kritikusok rámutatnak, hogy ez a keretezés túl leegyszerűsítő lehet (mivel az MI rendszerek magukban is beágyazhatják az elfogultságokat), de a könyv álláspontja egyértelmű: mi vagyunk felelősek az MI alkotásainkért. Ezért szinte minden történetben egy emberi főszereplő erkölcsi döntést hoz az MI-re reagálva – hangsúlyozva az ügynökségünket. Interjúkban Lee és Chen azt mondták, hogy a történetek azt akarták, hogy *"megerősítsék hitünket az emberi cselekvésben – hogy mi vagyunk a sorsunk urai"*.
- Társadalmi átalakulás és egyenlőtlenség: Számos vízió vizsgálja, hogyan súlyosbíthatja vagy enyhítheti a mesterséges intelligencia a társadalmi problémákat. Például az egyenlőtlenség visszatérő aggodalom – az 1. sztori a digitális kasztalapú diszkriminációval foglalkozik, a 8. és 10. sztori gazdasági megosztottságokkal foglalkozik. A szerzők elismerik, hogy az MI koncentrálna a vagyont és hatalmat (például nagy technológiai cégek, megfigyelő államok), de olyan ellenintézkedéseket is bemutatnak, mint az UBI, az adatszabályozás (GDPR a 9. történetben) és az oktatási reform (3. történet), hogy széles körben terjesszék az AI előnyeiket. A történetek globális környezete (India, Nigéria, Kína stb.) megerősíti, hogy az AI hatása világszerte fog elérni, nem csak a Szilícium-völgyben – és bizonyos esetekben a fejlődő országok is átugrhatnak vagy egyedi kihívásokkal szembesülhetnek.
- Ember-MI szimbiózis: Több fejezet (3, 5, 6) kiemeli az emberek és az MI közötti együttműködési kapcsolatokat. Ahelyett, hogy az MI egyszerűen helyettesítené az embereket, a könyv gyakran *ábrázolja a kiegészítést* – az MI oktatók fejlesztik a diákok

képességeit, MI társsátorok segítik az embereket, MI eszközök inspirálják a művészeket. Lee kifejezetten azt javasolja, hogy *"ember-AI szimbiózis" legyen, ahelyett, hogy azon töggednénk, mikor lesz az MI AGI-s.* A szerzők víziója szerint a legjobb eredmények akkor születnek, amikor az MI átveszi az irányítást ahhoz, amit a legjobban tud (adatfeldolgozás, rutinfeladatok), és felszabadítja az embereket, hogy arra koncentráljanak, amire a legjobban tudunk (kreatív, empatikus, stratégiai feladatok). Ez a téma az oka annak, hogy a könyv hangvétele nagyrészt reményteljes – az MI-t az emberi potenciál erősítőjeként ábrázolja, ha bölcsen kezelik.

- Etikai és egzisztenciális kockázatok: Bár összességében optimista, a könyv nem riad vissza attól, hogy *"mi tartja az AI vezetőit éjszakánként."* A 7. történet (kvantum/mesterséges intelligencia fegyverek) és a 2. történet (mélyhamisítással táplált káosz) éles figyelmeztetések arra, mi mehetne félre. Lee az MI átláthatósága, igazságosság és elszámoltathatóság kérdéseiről beszél – kiemelve, hogy a jelenlegi MI rendszerek lehetnek "fekete dobozok", amelyeket még az alkotók sem értenek teljesen. A közvélemény félelmeire is foglalkozik: az *"MI mint munkahely gyilkos"* félelem (ezért van szükség az újraképzésre és az UBI-ra), valamint az *"MI mint uralkodó"* félelem (azt állítja, hogy a szuperintelligencia szingularitása 2041-re valószínűtlen, de a szűk MI-t továbbra is óvatosan kell kezelni). A könyvből egy mondat foglalja össze: *"nagyon óvatosak azzal, hogy mit és hogyan optimalizálunk az MI-vel"* – mert a rosszul meghatározott célok katasztrófához vezethetnek (ahogy azt a Happiness Islanden vagy a biztosítási AI-nál láthatjuk). Végül soron a szerzők kiegyensúlyozott megközelítést javasolnak: fogadják el az AI lehetőségeit, de állítsanak be korlátokat (etikai kódexek, szabályozások, "MI jóért" kezdeményezések), hogy elkerüljük a disztópiát.
- A fikció találkozik a valósággal: A meta-téma a történetmesélés ereje a jövő alakításában. Chen Qiufan és Kai-Fu Lee szándékosan használják a fikciót arra, hogy az absztrakt jövők kézzelfoghatóvá váljanak, és vitát ösztönözzenek. Ez összhangban van a "sci-fi realizmus" hagyományával, amely befolyásolja a politikát (a könyv megemlíti, hogyan befolyásolták a futuristák a kormányzati gondolkodást Kínában). 2026-ra (ma is) már láthatjuk, hogy néhány *AI 2041* előrejelzés gyorsabban valósul meg a vártnál – például a ChatGPT és hasonló nyelvmodellek 2023-ra megjelentek, amelyek igazolják az AI oktatók "Iker veréb" vízióját; mélyhamisítások és autonóm járművek bizonyos mértékig jelenlegi valóságok. Valójában Lee és Chen forgatókönyvei konzervatívak lehetnek: *"20 éves idővonaluk meglehetősen [célszerűnek vagy akár több technológia esetében is] *konzervatívnak bizonyult,"* ahogy egy recenzív utólag megjegyezte. Ez aláhúzza a könyv lényegét, miszerint az MI forradalom most zajlik, nem valahol távoli jövőben, és fel kell készülnünk.

Fogadtatás és hatás

Az AI 2041-et széles körben jól fogadták, mint úttörő és könnyen érthető betekintést a közeljövőnkre. A 2021-es év legjobb könyvei közé választották olyan nagy médiumok, mint a *The Wall Street Journal*, a *The Washington Post* és a *Financial Times*. A *The New York Times* dicsérte, mint *"káprázatos... a jövőkép"* című könyvet írta, és a *FT* azt írta, hogy *"tele van érdekes meglátásokkal"*, a könyv borítójának szövegei szerint. Satya Nadella Microsoft vezérigazgatója dicsérte a kreatív történetmesélés és a technikai szakértelem ötvözését, mondván, hogy ez segít az olvasóknak *"megérteni, hogyan és mikor érnek bizonyos technológiák, és mit jelenthet ez mindannyiunk számára."* Az MI kiemelkedők, mint Yann LeCun (Turing-díjas) szintén támogatták a könyvet, és *"merész és sürgős"* együttműködésnek nevezte, amely betekintést nyújt abba, hogyan hat az AI az életünkre. Ezek a dicséret

kiemelik a könyv egyedi formátumát és oktatási értékét: sok olvasó számára lebilincselő módja volt az MI megismerésének, mert a fikció megismerheti az életet.

A kritikusok megjegyezték, hogy minden történet önmagában is meggyőző, és Chen Qiufan világépítési tehetsége ragyog. A globális környezetek és karakterek keveréke pozitív véleményeket kapott az AI hatását különböző kultúrákban és társadalmi-gazdasági kontextusokban való bemutatásáért, nem csak a Szilícium-völgyi nézőpontból. A történetmesélést Frank Diana technológiai stratégia "zseniálisnak" nevezte, aki azt mondta, hogy *"minden narratívában elmerült"*. A nagyközönség számára a könyv sikeresen tisztázza az MI-t: egy blogger *"gyorsúti tanfolyamnak nevezte az MI ernyője alatti technológiai fejlődéshez"*, és ajánlotta az MI műveltségének generációk közötti fejlesztésére. Valóban, az oktatók többgenerációs olvasási vitákra is fontolgatták.

Azonban nem minden visszajelzés kritikátlan. Néhány elemző rámutat, hogy Kai-Fu Lee, mint kiemelkedő MI befektető, teljes érdekű a mesterséges intelligencia által vezérelt jövő előmozdítása iránt, és a könyv technológia-optimistának vagy megoldásközpontinak tűnhet. A Los Angeles-i *Review of Books* *"Az AI 2041 semlegeségének zsarnoksága"* című kritikájában azt állította, hogy a mű *"egy olyan jövőt képzel, amelyben az MI a haladás központi hajtóereje"*, de *"nem vizsgálja eléggé"* az etikát, néha *"elhomályosítva"* a hatalmi dinamikákat. A LARB kritika azzal vádolta a könyvet, hogy a technológiát semlegesnek, a problémákat pedig több technológia által megoldhatónak ábrázolja, anélkül, hogy teljesen szembe kerüljenek azzal a lehetőséggel, hogy a technológia maga is elnyomó torzításokat vagy hatalmi egyensúlyhiányokat kódolhat. Például azt említi, hogy bár *az AI 2041* elismeri az algoritmikus torzítást (1. történet), az emberi adatokra hárítja a felelősséget, és nem vizsgálja, hogy az elfogultság *"elválaszthatatlan"* az MI tervezési döntéseiből. Emellett néhány olvasó úgy érezte, hogy néhány történet valamennyire didaktikus volt – a szereplők néha egy koncepciót illusztrálnak, ami a párbeszédet a leckéhez igazítva teheti. De ez talán elkerülhetetlen kompromisszum a könyv oktatási küldetése miatt.

A közönség szemszögéből a könyv számos vitákat indított el az MI-ről. Jó értékelést tart fenn az olvasóplatformokon (kb. 3,8/5 a Goodreads-en, a többség 4 vagy 5 csillagot ad). Az olvasók gyakran megjegyzik, melyik történet rezonált leginkább: az MI oktatók szívmengető látomása a "Ikerverebekben" egyesek kedvence, míg másokat a "Kvantumirtás" thriller vagy a "Boldogság szigete" filozófiai kérdései ragadtak meg. A formátumot még a *Black Mirror* epizódjaihoz is hasonlították, amelyeket magyarázó cikkekkel párosítottak – ez vonzó koncepció azok számára, akik szórakozást és elemzést is szeretnének. Másrészt néhány olvasó kissé sűrűnek vagy ismétlődőnek találta az elemző részeket, ha már eleve technológiailag jártasak, míg néhányan "túl optimistának" találták a fikciót. Egy BookWorm kritikusa megjegyezte, hogy a történetek *"szórakoztatóak... egy kissé disztópikus"*, de összességében a nagyközönségnek szól, így nem túl mély a technikai árnyalatokban.

A lényeg, hogy *az AI 2041* befolyásolta a döntéshozók és üzleti vezetők közötti beszélgetéseket is. Kai-Fu Lee fórumokon (például az USA-Kínai Kapcsolatok Nemzeti Bizottságának eseménye) beszélt a könyvből tanult tanulságokról, és annak forgatókönyveit arra használták, hogy szemléltessenék az MI irányítási keretrendszerek kidolgozásának sürgősségét. A konkrét narratívákat bemutatva absztrakciók helyett a könyv megkönnyíti a nem szakértők számára, hogy elképzeljenek, mi várható – esetleg alaposabb politikai vitákat eredményezve olyan témákban, mint az autonóm járművek szabályozása vagy a deepfake törvények.

Összefoglalva, az *AI 2041: Ten jövőképünk a jövőnk* átfogó képzeletgazdag, ugyanakkor reális feltárásaként mutatja be, hogyan alakíthatja át a mesterséges intelligencia életünk szinte minden területét a következő két évtizedben. A tíz fejezet-víziója – az MI-től az osztályterekben és kórházakban terjedő AI-ig a harctereken, gazdaságokon – együttesen azt állítja, hogy olyan átalakulások küszöbén állunk, amelyek egyszerre követelnek lelkesedést és éberséget. A szerzők egyedi együttműködésükön keresztül sikeresen oktatják és provokálják az olvasókat: gazdagabb megértéssel jutunk el olyan technológiákról, mint a mélytanulás, NLP, számítógépes látás, AR/VR, kvantumszámítás és még sok más – és ami még fontos, az emberi döntések meghatározzák, hogy ezek az eszközök disztópikus eredményekhez vagy egy virágzóbb, igazságosabb jövőhöz vezetnek-e. Ahogy Kai-Fu Lee írja, az MI "fordulóponton" van, és "*a lassú fejlődés napjai véget értek*". A könyv fő tanulsága a felhatalmazás: az AI pályáinak és társadalmi hatásainak előrejelzésével – egyénekként, közösségeiként és globális polgárokként – megalapozott lépéseket tehetünk, hogy a jövőt a kívánt víziók felé tereljük, nem csak azokhoz a célpont irányába, amelyek automatikusan megérkeznek.

Források: A fenti meglátások magukból az *AI 2041-ből* és hiteles bírálók elemzéséből származnak. Például a Penguin Random House összefoglalója tartalmazza a fejezet témáit és technológiákat, míg a Global Policy Journal és a Solventum technológiai blogjának részletes áttekintései a kulcsfontosságú történeti elemeket sorolják fel. A témák és a fogadtatás kommentárjai a LARB kritikáiból, szerzői interjúkból (pl. Diamandis blogja Kai-Fu Lee-vel), valamint a közvélemény reakcióiból származnak. Ezek a források együttesen gazdagítják az *AI 2041* képét, mint gondolkodási kísérletet és útmutatót – amely inspirálja a döntéshozókat (Satya Nadella "*magával ragadónak*" nevezte), mind az olvasókat arra, hogy átgondoltan foglalkozzanak a gyorsan közeledő MI jövőjével.