

Sfet Ingrid

Az erkölcsi autonómia kérdése a mesterséges intelligencia korszakában

Kivonat

A 21. század legújabb technológiai vívmánya a mesterséges intelligencia és az általa támogatott eszközök sora, melyek használata etikai aggályokat vet fel. Jelen tanulmány célja kiemelni és összegezni az említett technológia emberi autonómiára gyakorolt hatását, illetve feltérképezni ennek etikai vonzatát. Bemutatjuk és elemezzük azokat a mesterségesintelligencia-technológiákat és -modelleket, amelyek az egyén mindennapjaiban jelen vannak, annak érdekében, hogy pontosabb képet kapjunk azoknak az emberi autonómiára és erkölcsi cselekvésre gyakorolt hatásáról.

Abstract

The latest technological innovation of the 21st century is artificial intelligence, which is getting incorporated into more and more tools. The use and presence of these tools have brought attention to ethical issues; the aim of the present study is to highlight and compile the effect of the before mentioned technology on human autonomy and ethics. This study will mention different kinds of artificial intelligence technologies and models, which are present in our everyday lives, which will give a clearer picture on how these can affect human autonomy and ethical agency.

1. Bevezetés

„A mesterséges intelligencia etikájának kérdése annak meghatározása, hogy a mesterséges intelligencia hogyan segítheti elő az egyének jó életét vagy vethet fel azzal kapcsolatos aggályokat, akár az életszínvonalról vagy a demokratikus társadalmakhoz szükséges emberi autonómiáról és szabadságról van szó.”¹

¹ Európai Bizottság: *A megbízható mesterséges intelligenciára vonatkozó etikai iránymutatás*, 2018, B. I. fejezet, 32. cikkely, 11., <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (utolsó megtekintés dátuma: 2025. május 29.).

A 21. század a mesterséges intelligencia vívmányainak a korszaka, aminek a kezdeti fázisában már évek óta benne vagyunk. Ez nemcsak a társadalom kisebb rétegét érinti, hanem az átlagos ember mindennapjaiban is jelen van, anélkül, hogy az egyén tudna róla. Ennek a technológiai vívmánynak a jelentősége felér Neumann János informatika terén végzett munkásságával. Hagyatéka újabb mérőföldkőhöz ért, amikor a technológia arra a szintre lépett, hogy erkölcsi döntéseket képes hozni akár az emberek helyett is. Jelen tanulmány célja kiemelni és összegezni a mesterséges intelligencia emberi autonómiára gyakorolt hatását, illetve ennek etikai vonzatát. Bemutatjuk a legelterjedtebb MI-technológiákat és -modelleket, majd azt vizsgáljuk, hogy az új technológiák hogyan vetülnek az emberi autonómia erkölcsi síkjára. Ehhez szükséges tárgyalni az autonómia fogalomkörét és jelentését, az MI-technológiák működését, annak érdekében, hogy pontosabb képet kapjunk ezeknek az emberi autonómiára és erkölcsi cselekedetekre gyakorolt hatásáról.

Az autonómia nem más, mint egy ország, régió vagy szervezet önállósága, önálló szervezése és kormányzása annak érdekében, hogy függetlenül irányítsa magát, valamint az egyének vagy csoportok képessége arra, hogy cselekedjenek és döntéseket hozzanak anélkül, hogy bárki is irányítaná vagy befolyásolná őket.² Az autonómia a Merriam-Webster szótár³ szerint nem más, mint (1) az önálló irányítás állapota vagy minősége, pontosabban a saját kormányzás joga; (2) az önálló irányítási szabadság, avagy az erkölcsi függetlenség, ami az egyéni, személyes autonómiára utal; (3) egy önállóan kormányzó állam.

A tanulmánynak nem az a célja, hogy az egyéni autonómia fejlődését rekonstruálja, hanem az, hogy az egyéni autonómiát kontextusba helyezve világosabb képet mutasson be a mai modernizált világ erre gyakorolt befolyásáról. Az egyéni autonómia az évszázadok során jelentős változásokon és értelmezéseken esett át. Ma nyilvánvaló, hogy központi helyet foglal el a társadalometikában, és az is kézenfekvő, hogy előfeltétele a szabadság valamilyen formája. Az autonómia sajátosságai közé tartozik az, hogy méltóságot és jogot kölcsönöz az embereknek, mely kizárja, hogy eszközként kezeljék őket.⁴ Kant ezen nézete megalapozta az európai jogi és etikai nyilatkozatokat, mint például az *Emberi Jogok Egyetemes Nyilatkozatát*, mely már az első cikkelyében kiemeli, hogy „minden emberi lény szabadnak, egyenlő

² Oxford Advanced Learner's Dictionary: *autonomy* (szócikk). <https://www.oxfordlearnersdictionaries.com/definition/english/autonomy> (utolsó megtekintés dátuma: 2025. május 29.).

³ Merriam-Webster Dictionary: *autonomy* (szócikk). <https://www.merriam-webster.com/dictionary/autonomy> (utolsó megtekintés dátuma: 2025. május 29.).

⁴ Kant, Immanuel: *Groundwork of the Metaphysics of Morals*. (Eredeti cím: *Grundlegung Zur Metaphysik Der Sitten*, 1785.) Cambridge University Press, Cambridge, 1998. 54.

méltósággal és jogokkal születik”.⁵ Így az autonómia egyik szükséglete az emberi létezésnek, ennek korlátozása káros lehet mind az egyénre, mind a társadalomra nézve. A kérdés abban rejlik, hogy a mai modernizált világ technológiája miként befolyásolja vagy befolyásolhatja az egyén személyes autonómiáját, valamint abban, hogy mit tehetünk a káros befolyások ellen.

2. Az autonómia fogalomköre

Amint láthattuk, az autonómia fogalma, kontextustól függően, több jelentést is kaphat. Ez jelölheti az államok politikai függetlenségét, az egyén és az intézmények önállóságát is. Dietz szócikkében egy pontosabb meghatározást ajánl az autonómia kapcsán. Elsősorban kiemeli a kifejezés forrását, miszerint az autonómia az *αὐτόνομος* görög szóból ered, illetve hogy Hérodotosz értelmezésében a „saját akaratunk szerint viselkedni” jelentése van, a politikai függetlenséget hangsúlyozva.⁶

A fogalom rekonstruálása érdekében érdemes kiemelni, hogy az autonómia kifejezés a középkorig feledésbe merült, majd csak 1586-ban vette fel újra Francisus Burgkardus *De autonomia* című művében, ahol az „autonómia” valláspolitikai jelentése a „diszkrécio” a felekezet megválasztásában.⁷ A reformáció egy felekezeti megosztottságot vont maga után, amelynek egyik eredménye az egyének szabadságáért folytatott viták felerősödése volt. Melanchthon a keresztény szabadságot a vallásos egyén alapjogának tartotta. A szabadságról folytatott viták egyik eredménye az egyházi tekintélyekkel szemben megerősödött *libertas christiana*⁸ és a világi rend viszonylagos autonómiája. Megfigyelhető, hogy a wittenbergi reformátorok, az anabaptista csoportok, a reformációt támogató császári birtokok és a svájci reformáció teológusai között voltak nézeteltérések, amelyek magukba foglaltak jogi vitákat is a *libertas religionis*, a *libertas conscientiae* és az *autonomia* körül.⁹ Ez jelentős, mivel a vallás és a lelkiismeret szabadsága együtt jártak. Ezek csak az 1949-es német alaptörvényben¹⁰ különültek el. A vallásszabadság az egyén szabad vallásválasztását és -gyakorlását védi. A lelkiismereti szabadsággal az egyén lelki identitá-

⁵ ENSZ: Emberi Jogok Egyetemes Nyilatkozata, 1948. december 10., 1. cikk.

⁶ Lange, Dietz: Autonomy (szócikk). In: Dietz Lange (szerk.): *Religion Past and Present Online*. Brill, [h. n.] 2011.

⁷ Uo.

⁸ Keresztény szabadság.

⁹ Kaiser, Otto et al.: Freedom (szócikk). In: Dietz Lange (szerk.): *Religion Past and Present Online*. Brill, [h. n.] 2011.

¹⁰ Lásd A Német Szövetségi Köztársaság alaptörvénye, 4 art. 1–3. cikkely, 1949. 05. 23.

sát védték, hisz ez kivételesen felmentheti őt a törvény által előírt cselekvések és az abszolútnak tekintett belső kötelezettségek alól.¹¹ Ez a kiváltság védi az egyén azon jogát, hogy saját lelkiismerete szerint hozza meg az erkölcsi döntéseit.

Teológiai szempontból látható, hogy 1780 óta vitatkoznak a formálisan elfogadott negatív szabadságról és pozitív szabadságról.¹² A 18. századi felvilágosodás, az idealizmus és a német historizmus teológusai a vallási szövegeknek azon oldalát emelték ki, amelyek a szabadságot hangsúlyozták, ezúton az emberi és társadalmi jogokat a kereszténységre vezették vissza.¹³ A protestáns teológusoknak a reformáció volt a központi kiindulópontja. Hegel ennek tartalmát így határozza meg: „ez a reformáció lényege: az ember önmaga határozza meg, hogy szabad legyen”.¹⁴ Így a 18. századtól a 20. századig számos protestáns kutató igazolta a modern ember autonómiáját mint a „protestantizmus” elvét. A szabadság igen vitatott fogalom volt, mivel egyrészt a különböző csoportok változó értékekhez kötötték, másrészt pedig az egyház intézményes berendezkedése megnehezítette a szabadság vallási és politikai jelentésének az elkülönítését – magyarázza Friedrich Wilhelm Graf.¹⁵ Így előtérbe került a negatívan és pozitívan meghatározott szabadság, ahol a pozitív szabadság a vallás birtoklását és gyakorlását védi, míg az említett alaptörvény a negatív szabadságot képviseli és védi, ami a kényszerített vallásgyakorlást és birtoklást tiltotta.¹⁶ Itt az egyén a lelkiismeretének kell hogy engedelmességedjen minden erkölcsi döntés esetén. Felmerül a kérdés, hogy létezik-e moralitás vallásosság nélkül.

2.1. Autonómia mint a moralitás és vallásosság kérdésköre

Fairweather és McDonald tanulmányukban¹⁷ a vallás és moralitás összefüggését vizsgálják négy szempontból. (1) *A moralitás függ a vallástól*, avagy a moralitás nem létezik vallás nélkül, mivel ez adja meg a jó és rossz közti különbségtétel alapját, aminek visszaigazolása lehet a vallás, ami csakis Isten által lehetséges. (2) *A vallás függ a moralitástól* – Braithwaite nézeteire alapozott tétel, ami szerint

¹¹ De Wall, Heinrich: Freedom of Belief (szócikk). In: Dietz Lange (szerk.): *Religion Past and Present Online*. Brill, [h. n.] 2011.

¹² Kaiser et al.: i. m.

¹³ Uo.

¹⁴ Hegel, G. W. Friedrich: *Vorlesungen über die Philosophie der Geschichte, Theorie Werkausgabe XII*. Reclam Verlag, [h. n.], 1970. 496.

¹⁵ Kaiser et al.: i. m.

¹⁶ De Wall: i. m.

¹⁷ Lásd Fairweather, Ian C. M.: *The quest for Christian ethics: An inquiry into ethics and Christian ethics*. Edinburgh, Handsel Press, Edinburgh, 1984. 95.

az erkölcsösség és moralitás a hagyományokból merítenek. Ezt a szerzők vitatják, mivel a vallások komplexumában az erkölcs csupán egy tényező a sokból. (3) *A moralitás és vallás különböző valóság, egymástól független szféra* – ez annak tudható be, hogy az erkölcs racionális akaratból származik, ami szembenmegy a vallással, mivel ez kétségbe vonja az erkölcs autonómiáját. Ez azonban nem jelenti azt, hogy ne lenne más kapcsolat köztük. (4) *A moralitás és vallás különböző, de nem függetlenek* – a kettő különbözik, viszont egymástól függenek, saját autonómiájuk van, de mégis kapcsolatban állnak.¹⁸ Konklúziójukban kiemelik, hogy a négy pontban felsorolt nézeteltérések tárgyalása szükséges volt több aspektusból is, és feltételezhető, hogy a vallás és moralitás közti kapcsolat egyedi, mindkettő rendelkezik a maga autonómiájával, és egyiket sem kell a másakra redukálni, mégis belsőleg kapcsolódnak egymáshoz. Ez a kapcsolat nem formális és külső megnyilvánulásként, hanem a belső életünkben van jelen.¹⁹

Németh Dávid *Vallásdidaktika* című kötetében négy érvet emel ki, miszerint a hit- és erkölcsstannak helye van a közoktatásban, majd etikai érvében magyarázza, hogy a társadalmi kapcsolatok működőképességét egyedül a jogrendszer nem képes fenntartani: az igazi összetartó erőt a moralitás adja.²⁰ Kiemeli, hogy a gyermekek nevelésében az erkölcsi nevelés egy kulcsfontosságú szerepet játszik, mivel enélkül az általános moralitás színvonala fokozatosan romlik a társadalomban, és a családok ezt egyre kevésbé tudják betölteni. Az oktatás mellett ez a gondolat rávehető a közösségekre és a társadalom egészére is, mivel az emberi identitás egyik jellemzője a folyamatos fejlődés.

Az autonómia kifejezést gyakran használják abban a kontextusban, hogy nem valamilyen külső tekintély kényszeríti ki, hanem a racionális akarat maga szabja meg. Ez az erkölcsi fejlődés végső szakaszának a jele, amikor a felnőtt saját akaratából fogadja el a normákat, és nem külső hatás befolyására teszi ezt. Robinson megfogalmazta, hogy az erkölcsiség természeténél fogva magába foglal egy törvényt, amely független a cselekvő egyén vágyaitól. Ezzel kapcsolatban felmerül az autonómia és a heteronómia kérdésköre, mivel ezt kétféleképpen is fel lehet fogni. Vagy magunk szabjuk meg a normákat, vagy úgy fogjuk fel, mint egy kívülről ránk kényszerített tényezőt. Autonómia alatt tehát az erkölcsi életnek olyan felfogása értendő, ami egy olyan törvény szerint él, ami belülről ered.²¹ Így látható, hogy az

¹⁸ Uo. 100–112.

¹⁹ Uo. 111–112.

²⁰ Németh Dávid: *Vallásdidaktika. A hit- és erkölcsstan tanítása az 5–12. osztályban*. L'Harmattan Kiadó, Budapest, 2017. 27–29.

²¹ Robinson, N. H. Galloway: *The Groundwork of Christian Ethics*. Collins, United Kingdom,

egyéni autonómia belülről fakad, és nem külső tényezők összessége. Tehát előtérbe kerülnek az egyén személyes döntései, értékei, tapasztalatai, a szocializáció során áthagyományozódott vagy elsajátított erkölcsi minták és meggyőződések.

2.2. Az autonómia és az erkölcsi meggyőződés, cselekvés

Heller Ágnes *Általános etika* c. kötetében magyarázza a cselekvés moralitását, szándékosságát és szándéktalanságát. Az erkölcsi megítélés mindig felveti a cselekedetek szándékosságának vagy szándéktalanságának és az egyén szabad választásának a kérdését. Heller Ágnes számára adott esetben fontos eldönteni, hogy mennyiben tulajdonítható valamely cselekedet egy személynek mint erkölcsi cselekvőnek, mivel ahol cselekvés van, ott van felelősség is.²² Viszont a felelősség jellege és foka már nem a cselekvés szándékosságától függ, hanem sok más tényezőtől. Ebben az esetben Heller Ágnes a szabad választás elméletét emeli ki, amikor az egyénnek lehetséges cselekedetek közül kell választania, és a választáshoz vezető út csak akkor nyílik meg, amikor az egyén a helyzetét egy erkölcsi helyzetnek tekinti, így előtérbe került, az erkölcsi megfontolás és megítélés.²³ Így felmerül a kérdés, hogy egy egyén lehet-e teljesen autonóm. Erre a kérdésre keresve a választ Heller Ágnes a következőképpen fogalmazott:

„...az abszolút autonómia azt jelenti, hogy a személy mint személy teljesen szabad. A viszonylagos autonómia azt jelenti, hogy a személy mint személy bizonyos fokig szabad, de nem teljesen az. Az abszolút heteronómia azt jelenti, hogy a személyt mint személyt teljesen rajta kívül álló tényezők határozzák meg. A viszonylagos heteronómia azt jelenti, hogy a személy mint személy bizonyos, rajta kívül álló kényszerek vagy hatalmaknak van aláéve. A viszonylagos heteronómia előfeltételezi a viszonylagos autonómiát és viszont, míg az abszolút autonómia az abszolút és viszonylagos heteronómiát egyaránt kizárja. Az abszolút autonómia azt jelenti, hogy a személy teljesen szabad erkölcsi alany.”²⁴

Így látható, hogy az abszolút autonóm erkölcsi személy csak egy olyan egyén lehet, aki olyan normáknak és szabályoknak megfelelően cselekedhet, amelyeknek érvényességét a lelkiismeretéhez köti, és semmilyen kényszert nem érez arra, hogy bármely normát vagy szabályt megszegjen. Ezzel ellentétben találkozunk az ab-

1972. 172.

²² Heller Ágnes: *Általános etika*. Cserépfalvi Könyvkiadó, Szolnok, 1994. 66.

²³ Uo. 74–75.

²⁴ Uo. 77–78.

szólút erkölcsi heteronómiával, mikor egy személy bábu módjára viselkedik minden erkölcsileg jelentős cselekedet alkalmával. Ezt a befolyást kétféle tényező is meghatározhatja: környezetének normái és szabályai, valamint személyes érdekei. Ebből a pontból felépíthető a viszonylagos erkölcsi heteronómia és a viszonylagos erkölcsi autonómia is. Az autonómia valamennyi lehetséges cselekedetre vonatkozik, míg az erkölcsi autonómia csak az erkölcsi jelentőségű vagy tartalmú cselekedetekre, és ez érvényes a heteronómiára is. Viszont felmerül a kérdés, hogy mi előzi meg a cselekedeteket, legyen az autonóm vagy heteronóm jellegű.

Szűcs Ferenc *Teológiai etika* c. könyvében magyarázza, hogy maga az etika mint az erkölccsel foglalkozó tudomány már az ókortól vizsgálta, hogy milyen emberi cselekedetek tekinthetők erkölcsösnek, igazságosnak, valamint azt, hogy milyen szempontok alapján ítéltük meg azokat jónak vagy rossznak. Szűcs Ferenc egy hármas rendszert emelt ki, amiben felvetődik a jó és a rossz kérdése: indíték, cél és a kiváltott hatás.²⁵ Az indíték az emberi akarathoz kapcsolódik, vagyis ahhoz, hogy az egyén képes dönteni, így ez megelőzi a cselekvést. Ezért már az indítéknál felvetődik a jó és a rossz kérdése. A cél feltételezi a racionalitást, tehát értelmet ad a cselekvésnek, valamint feltételezi, hogy a személy képes előre látni cselekvésének a célját. Így a folyamatban jelen van a megítélés és a meggondolás. A kiváltott hatás eltér az előzőktől, mivel ellentmondásos is lehet. A jó akarat és szándék nem mindig igazolja a cselekedetek helyességét, mivel a tréfás, szórakoztató cselekedeteknek is lehet negatív kivetülése. Ebben az esetben megmutatkozik a jó és a rossz azon szubjektív megítélése, amely viszonylagossá teszi az etikát, ezért szükség van konszenzuson alapuló normákra.²⁶

A szerző továbbá elemzi a természeti törvény azon egyoldalú értelmezéseit, amelyből született az autonóm erkölcsiség fogalma.²⁷ Kiemeli, hogy az autonómia nem tudott választ adni az erkölcs tartalmi kérdéseire, és „a jó szubjektív megítélése maradt a gyöngé pontja”.²⁸

Fazakas Sándor *A protestáns etika feladata, alapjai, hagyományai és az erkölcsi meggyőződés kialakítása* c. tanulmányában²⁹ tárgyalja az erkölcsi meggyőződés

²⁵ Szűcs Ferenc: *Teológiai etika*. A Református Zsinati Iroda Tanulmányi Osztálya, Budapest, 1993. 2.

²⁶ Uo. 2–3.

²⁷ Uo. 7.

²⁸ Uo.

²⁹ Lásd Fazakas Sándor: *A protestáns etika feladata, alapjai, hagyományai és az erkölcsi meggyőződés kialakítása*. In: Uő (szerk.): *A Protestáns etika kézikönyve*. Kálvin János & Luther Kiadó, Budapest, 2017. 15–49.

kialakításának a lépéseit. Az erkölcsi meggyőződések és döntések fontos szerepet játszanak a lehetséges cselekvések mérlegelésében. A 21. századi ember olyan kihívásokkal és veszélyekkel áll szemben, amelyekre megoldásokat kell találnia és felelősségteljes döntéseket kell hoznia. Ennek érdekében elengedhetetlen az erkölcsi meggyőződések kialakításának, az ítéletalkotásnak a képessége.³⁰ Az erkölcsi véleményformálás nem csak egy racionális folyamat, mivel vannak olyan helyzetek, amikor egy döntést gyorsan kell meghozni. Az ilyen döntések egy másik indítéka az intuíció, amely számára meghatározóak az emlékek, élmények, tapasztalatok és egyéb szocializáció során kialakított, áthagyományozódott vagy elsajátított cselekvési minták. Ehhez úgyszintén hozzátartoznak az egyén tapasztalata vagy neveltetése során kialakított erkölcsi meggyőződései is.³¹ Ebből kifolyólag látható, hogy a cselekedeteket megelőzi az erkölcsi meggyőződések kialakítása, az erkölcsi véleményformálás és az intuíció.

2.3. Az autonómia fogalomkörének aktualitása

Az előzőekben tárgyaltuk az autonómia fogalomkörét, és nyilvánvalóvá vált ennek komplexitása. Megvizsgáltuk azt, hogy mit foglal magába az autonómia, valamint azt is, hogy az embernek csak viszonylagos autonómiája és viszonylagos erkölcsi autonómiája lehet. Így felmerül a kérdés, hogy az aktuális, MI által támogatott digitalizált világban ez miként tükröződik vissza. Carina Prunkl ezt úgy közelítette meg, hogy az emberi autonómiát két oldalról elemezte: belső és külső dimenzióból.³²

A belső dimenzió a személy azon képessége, hogy saját értékei és meggyőződései szerint döntsön és cselekedjen. Ez feltételezi a külső hatások (jelen esetben technológiai hatások) értékelését, azt, hogy ezek nem hatnak az egyénre manipulatív és megtévesztő módon. Ebben az esetben az autonómiának szüksége van a racionális gondolkodásra, önreflexióra, önkontrollra és a megfelelő információk megfelelő alkalmazására.³³ Ezzel együtt a szerző azt is kiemeli, hogy minden egyéniség bizonyos mértékben élményeinek, tapasztalatainak és közösségi hagyományainak értékeit hordozza önmagában.³⁴ Ez a gondolat párhuzamba állítható az erkölcsi döntések intuitív indítékaival.³⁵

³⁰ Uo. 36.

³¹ Uo. 46–47.

³² Lásd Prunkl, Carina: Human Autonomy at Risk? An Analysis of the Challenges from AI. *Minds and Machines*, 2024. 34:26.

³³ Uo. 8.

³⁴ Uo.

³⁵ Lásd Fazakas: *A protestáns etika kézikönyve*. 46–47.

A külső dimenzió az erkölcsi motivációk gyakorlati aspektusára fókuszál, ami magába foglalja azt, hogy olyan értelmes választási lehetőségek állnak rendelkezésünkre, amelyek mentén az egyén általánosan tudja irányítani az életét. Ez feltételezi a szabadságfeltételeket mint vallási és politikai szabadságot, valamint a lehetőségfeltételeket, amely feltételezi a lehetőségekhez való megfelelő hozzáférést. Ezek a feltételek kívül esnek az egyén hatáskörén, és jelentős mértékben függ az egyén társadalmi hovatartozásától, valamint a jogi, gazdasági és politikai infrastruktúrától.³⁶ Carina Prunkl megközelítése itt újból előtérbe helyezi az autonómia és heteronómia kérdését, miszerint az egyén dönthet és cselekedhet autonómiájának megfelelően, vagy ellenkező esetben „áldozata” környezetének, és nem tudja érvényesíteni személyes erkölcsi értékeit.

3. Egyéni autonómia az MI korszakában

A jelenlegi technológiai igyekezetek sok tekintetben előnyt jelentenek a felhasználók körében, viszont nem szabad figyelmen kívül hagyni a velük járó hátrányokat és kockázatokat. A technológiák megváltoztatják az egyének interakcióját a társadalommal és a világgal, és így a legújabb technológiák, mint az MI, módosítják az emberi cselekvés feltételeit és lehetőségeit.

Nyilvánvaló, hogy a legtöbb MI-technológia Machine Learning (ML), avagy gépi tanulási folyamatokat alkalmaz. A gépi tanulási folyamatok egyik köztudott algoritmusa a Deep Learning (DL), avagy mély tanulási funkció. Míg az MI az emberi viselkedést és intelligenciát egyesíti gépekben és/vagy rendszerekben, addig az ML adatokból és/vagy tapasztalatokból tanul, annak érdekében, hogy automatizálja fejlődését. A mélytanulás (DL) szintén egy adatvezérelt tanulási módszer, amely az ember agyműködését utánozza. Ezek mind olyan alapvető MI technológiák, amelyek az intelligens rendszerek fejlesztésére és a folyamatok automatizálására használatosak.

Így a következőkben az MI-technológiák társadalom- és szociális vonatkozásait tárgyalom, a teljesség igénye nélkül, viszont kiemelve a téma kutatásának és feltárásának a szükségességét. Jelen fejezetnek nem célja az MI-technológiák működésének és fejlődésének a feldolgozása, viszont annak érdekében, hogy világos képet kaphassunk arról, hogy az MI milyen kölcsönhatásban van az egyénnel és a társadalommal, fontos működésének az átláthatósága és felhasználási céljainak a tisztázása.

³⁶ Prunkl: i. m. 11.

3.1. A generatív MI etikai kockázatai

A generatív MI egyik kézenfekvő és közismert példája a ChatGPT. A generatív MI-modell alkalmazása számos lehetőséget nyújt tartalmak létrehozására, legyen az szöveg, kép, videó, hang, beszéd, zene, szoftver kód, design vagy művészeti alkotások, így számos területen számos célra alkalmazható. Viszont kihívásokkal, kockázatokkal jár úgy fejlesztői, mint felhasználói szempontból: az MI-hallucinációk és a képzési adatok mai napig kihívást jelentenek a fejlesztőknek, ugyanígy a *deepfake* tartalmak és a biztonságot, magánéletet vagy szellemi tulajdont fenyegető veszélyek, melyek az MI negatív felhasználási példái közé sorolhatók. Ezek a kihívások és kockázatok sok esetben olyan hibák, melyek az algoritmus pontatlanságának az eredményei, és még fejlesztést igényelnek. Felhasználói szempontból pedig olyan kihívásokkal és kockázatokkal találkozunk, amelyek kifogásolható erkölcsi meglátások és cselekedetek miatt születtek. A generatív MI-modell esetében tárgyaljuk a fejlesztői kihívások irányából érkező aktuális etikai kockázatokat, a hallucinációkat és a képzési adatok minőségét, majd a felhasználói oldalról érkező kihívásokat is elemezzük.

MI-hallucinációk és a képzési adatok minősége

A hallucinációk olyan MI-kimenetek, amelyek értelmetlen vagy pontatlan eredményeket adnak felhasználójának. A gyakorlati szakemberek ezt elkerülhetetlen következménynek tekintik a modell pontosságából és kreatív képességéből adódóan.³⁷ Ebben az esetben a képzési adatok minősége is egy olyan kihívás, amellyel a generatív MI szembesül, mivel ennek kimenetelei nagyban függenek a beletáplált adatok minőségétől.³⁸ Ennek érdekében a fejlesztők azon dolgoznak, hogy a modell releváns és megbízható adatokból táplálkozzon, valamint a folyamatos értékelés is segíthet a pontatlan eredmények csökkentésén.

A felhasználó kritikus gondolkodása itt kulcsfontosságú szerepet kap. Az említett hallucinációk, a felhasznált területtől függően, befolyásolhatják az autonóm egyén döntéseit, cselekedeteit, értékeit, sőt esetenként gondolatait is. A kritikus gondolkodás az MI felhasználása során fontos tényező a nem hiteles vagy hibás kimeneteli eredmények felismerésének esetében.

³⁷ Stryker, Cole – Scapicchio, Mark: *What is generative AI?* <https://www.ibm.com/think/topics/generative-ai> (utolsó megtekintés dátuma: 2025. május 29.).

³⁸ Fui-Hoon Nah, Fiona et al.: *Generative AI and ChatGPT: Applications, Challenges, and AI-Human Collaboration. Journal of Information Technology Case and Application Research*, 2023/25 (3). 277–304. 288.

Ajánlórendszerek

Az MI által alkalmazott ajánlórendszerek sokasága és komplex alkalmazása a digitalizált eszközök számos területén jelen van. Ezekkel találkozunk a közösségi médiában, hírportálokon és egyéb digitális közösségi tereken. Ezt egy burkolt manipulációnak is lehet nevezni, ahol az adaptív tartalomajánlás elemzi felhasználójának viselkedését, és erre a viselkedésre építve különböző személyre szabott ajánlatokkal látja el.³⁹ Ez a jelenség manipulatív módon befolyásolja a felhasználó döntéseit, nézeteit és értékeit. Egyes felhasználók pedig hajlamosak arra, hogy az adott technológiai rendszer által adott javaslatot részesítse előnyben. Ebben az esetben az MI befolyásolhatja az egyén azon képességét is, hogy megfelelő ítéletet hozzon. Ilyen esetekkel több szinten találkozhatunk, mint például egy árucikk megvásárlásának a döntése vagy az összeesküvést szító tartalmak felerősödése, amely ösztönzője lehet a hamis hiedelmek, értékek és ítéletek kialakításának.

Hasonló példa az emberi értékek eltorzulása, mivel a közösségi médiában megjelenő tartalmak a teljes toleranciát képviselik, amivel olyan dolgok válnak elfogadottá, amelyek károsak az ember személyes életére vagy akár egészségére, identitására nézve is. Manapság az öndiagnózis virágzik, ezért szinte „trendi” az, ha valamilyen egyedi tulajdonsággal kitűnhetünk a tömegből. Az MI által támogatott marketingalgoritmusok azokat a profilokat támogatják, amelyek valami egyedit tudnak nyújtani a felhasználóknak. Ez heteronóm megfelelési kényszer által vezérelt autonóm döntésnek is nevezhető, mivel az öndiagnózison keresztül egy kisebb közösséghez tartozhatnak az egyének, így kitűnve a tömegből. Ilyen megosztó tartalom lehet az ADHD, kövérphobia, homoszexualitás, gender-spektrum stb. Ezek sok esetben éppen az ellenkezőjét érik el annak, ami ténylegesen fontos: nem az elfogadást és a segítségnyújtást ösztönzik.

Az egyéni autonómiát veszélyeztető kockázatok elkerülése érdekében több tényezőt kell figyelembe vennünk. Ilyen esetekben újból előtérbe kerül a kritikus gondolkodás mint alapvető kompetencia, az alapvető technológiai ismeretek elsajátítása és alkalmazása, valamint az erkölcsi ítéletalkotás fontossága. Ezzel párhuzamosan nyilvánvaló, hogy figyelembe kell venni az egyén társadalmi integrációját, társadalmi és közösségi értékeit, oktatási és neveltetési háttérét és egyéb olyan tényezőket, melyek meghatározzák az egyén intuícióját, amelyek hatással lehetnek az erkölcsi ítéletalkotására.

³⁹ Stryker – Scapicchio: i. m.

Deepfake tartalmak

A *deepfake*-ek olyan MI által generált vagy manipulált képek, videók vagy hangok, amelyek a felhasználó számára valóságnak mutatnak olyan megnyilvánulásokat, amelyeket a képen látott személy vagy csoport sohasem mondott vagy tett. Ezek mind olyan jelenségek, amelyeket a felhasználók rosszindulatú célokkal generálnak. A kutatók dolgoznak olyan MI-n, amelyik képes felismerni a *deepfake*-et, de addig a felhasználók oktatása és képzése a legjobb eljárás az átverések elkerülése érdekében.⁴⁰ Ezekkel a jelenségekkel tudat alatt is befolyásolhatóak az egyén erkölcsi és általános autonóm értékei, döntései, motivációi és meggyőződései.

A magánéletre is veszélyesek a *deepfake* jellegű termékek, amelyeket zaklatás, cyber-bullying céljával gyártanak, esetenként humoros jelleggel. Több oktatási intézményben találkozhatunk zaklatással gyerekek között, amit az MI tesz hangsúlyosabbá, mivel a diákok saját társaikról pillanatok alatt kompromittáló anyagokat tudnak létrehozni. A valóban autonóm egyénnek lehetősége van, hogy az említett szoftvereket saját döntései szerint használja vagy mellőzze. Az MI ebben az esetben eszköz csupán – kérdés, hogy a felhasználó autonóm egyénként milyen célokra veszi igénybe ezt a technológiát.

A generatív MI esetében fontos kiemelni a nyelvi modelleket is, amelyek használhatóak adathalász e-mailek, hamis személyazonosságok vagy más rosszindulatú tartalmak létrehozására. Ezek olyan felhasználói alkalmazások, melyekkel a résztvevők nemcsak manipuláció és átverés áldozataivá válnak, hanem megszerzett adataik révén személyi integritásuk is kompromittálható. A fejlesztőknek és a felhasználóknak egyaránt ügyelni kell arra, hogy a modellbe betáplált adatok ne tegyék hozzáférhetővé a saját szellemi tulajdonukat vagy más szervezet szellemi tulajdonát.⁴¹

Az MI erősítheti ugyan a szabad információáramlást és az újságírói tevékenységet, ám ezzel együtt a hamis nyereményjátékok, álhírek és internetes átverések is jelen vannak. Itt az egyén autonómiája több irányból érintett. A tudatalatti manipuláció az egyik közismert kihasználása az említett MI nyelvi modelleknek. A választópolgár politikai elfogultságának a megelőzésére tervezett algoritmusok eldöntik, hogy milyen tartalmak jelenjenek meg a közösségi médiában. Ettől továbbhaladva ugyanez az algoritmus kihasználható arra is, hogy hamisított, manipulatív tartalmak terjedjenek el a közösségi médiában különböző célcsoportok számára. Ezek lehetnek érzelmi propagandák is. Itt az autonómia szempontjából

⁴⁰ Uo.

⁴¹ Uo.

a manipuláció és az átverés készíti az egyént nem autonóm cselekedetekre, vagy ellenkező esetben a nem cselekvésre vagy autonóm véleményének a kinyilvánítására információ hiányában. Így ez hatással van a polgári tájékozódás és a társadalmi bizalom normáira is.

3.2. A természetes nyelvfeldolgozás (NLP) etikai kockázatai

A Natural Language Processing (NLP), avagy a természetes nyelvfeldolgozás az MI olyan részterülete, amely a gépi tanulást használja arra, hogy a számítógépek képesek legyenek megérteni az emberi nyelvet, és kommunikálhassanak vele. Ez lehetővé teszi a szöveg és beszéd felismerését, megértését és generálását, a statisztikai modellezést, ML és a mélytanulás kombinálásával. Az NLP lehetővé tette a nyelvi tanulási modellek (LLM) kommunikációs képességeitől kezdve a komplexebb feladatok végzését is, mint a kép, hang és videó generálását. Ez már a mindennapi élet része, mivel hajtja a keresőmotorokat, ügyfélszolgálati chatbotokat, hangvezérelt GPS-eket, okostelefonok asszisztenseit.⁴²

Ezzel egyidejűleg az NLP-nek is vannak kihívásai. Az emberi nyelv többretegű, így a programozók számára is nehézséget jelent olyan szoftvert írni, amely pontosan meghatározza a szöveg vagy hangadatok szándékolt jelentését. A programozóknak meg kell tanítaniuk a természetes nyelven működő alkalmazásokat a szabálytalanságok felismerésére és megértésére, hogy alkalmazásuk pontos legyen. A nehézség a nyelv természetes „dallamából” adódik, amire a természetes nyelv kifejezés is utal, mivel ezeknek a modelleknek a célja, hogy a szöveget, hangot és képet minél emberibb módon tudja reprodukálni, ami nagyban függ a kiválasztott nyelvtől, azon belül a nyelvi használók régiójától, akár a reprodukált személy sajátos intonációjától is. Az emberi kommunikáció egy másik aspektusa a hangnem, amikor a testbeszéd és a szóbeli előadásmód teljesen más jelentést ad a mondánivalónak, mint a szavak használata. A hatáskeltő túlzások, szavak hangsúlyozása, szarkazmus miatt az NLP összekeverheti őket, ami kevésbé megbízhatóvá teszi a szemantikai elemzést.⁴³ Viszont itt is találkozunk kockázatokkal: ilyen lehet a hangnem helytelen értelmezése, a félreértés lehetősége, az előítéletek hatása vagy az egyéni és társadalmi szabadságot fenyegető alkalmazások lehetősége.

Felhasználói szempontból a természetes nyelvfeldolgozás technológiája elérhető és alkalmazható a generatív MI-modellek esetében, ahol az álhírek, manipulá-

⁴² Stryker, Cole – Holdsworth, Jim: What is NLP (natural language processing)? <https://www.ibm.com/think/topics/natural-language-processing> (utolsó megtekintés dátuma: 2025. május 29.).

⁴³ Uo.

ciós tartalmak újból megjelennek. Az erkölcsileg megkérdőjelezhető tartalomgyártások esetében találkozunk olyan anyagokkal, melyben egy társadalmi képviselő és/vagy közismert egyén videójára, hangjára egy olyan hanganyagot vágnak rá, amellyel összeesküvést szító, társadalmi manipulációs tartalmakat osztanak meg. Az ilyen *deepfake* jellegű tartalmakat a generatív MI-modellek esetében tárgyaltuk, az egyéni autonómiára gyakorolt kockázatokra utalva.

Gépi fordítások

Az NLP-nek köszönhetően előtérbe kerültek a gépi fordítások (Machine Translation). Ezek a fordítások egyre elterjedtebbé válnak úgy a fordítói szakmán belül, mint a felhasználók körében. A gépi fordítás sok esetben előnyt és gyorsaságot jelent felhasználójának, viszont sok esetben nyelvi hibákkal is találkozunk: mondatfelépítés, hangnem helytelen használata, mondatkiegészítés. Így a gépi fordítás mondatstruktúrája befolyásolja a felhasználó eredeti mondatfelépítését, szintaxisát és eredeti mondanivalóját. Sok esetben ez nem koherens fordításhoz is vezethet, mivel mondatonként viszi be, és így a forrásnyelvben használt egységes kifejezések esetében a célnyelvben nem egységes fordítással találkozunk, hanem ezek szinonimáival, amelyek árnyalati eltéréseket eredményeznek.

Hasonló tartalmakkal és szövegekkel több környezetben is találkozhatunk, ahol a gépi fordítás nem megfelelő termékleírással vagy tartalommal látja el felhasználóját. Az autonómiára gyakorolt kockázat esetében itt már nem csak a szöveg értelemzéséről van szó, hanem a felhasználóra gyakorolt hatásról is. Amennyiben egy idegen forrásnyelvből szükséges dolgoznia egy felhasználónak, és a tartalom megértése és munkája gyorsítása érdekében gépi fordítást alkalmaz, majd azzal dolgozik hosszabb ideig, akkor ez torzult nyelvhasználathoz és kommunikációhoz vezethet úgy szakmai tevékenységében, mint mindennapjaiban. A generatív MI-hez hasonlóan a gépi fordítások is képzési adatokból és betáplált információkkal dolgoznak, amelyeket nem minden esetben tudnak megfelelően párosítani, így a kimeneteli fordítások esetében újból a felhasználó kritikus hozzáállására és ellenőrzésére van szükség.

Felhasználói lehetőséglimitálások

A programozásnál itt is fennáll a *garbage in, garbage out* (GIGO) veszélye, ami nem más, mint a betáplált adatok minősége. A beszédfelismerés egyik problémája a félreértelmezés, ahol az NLP-megoldások összezavarodhatnak, ha a beszélt bemeneti adatok homályos dialektusban, motyogva, szlenggel, helytelen nyelvtannal,

félreérthető kiejtésekkel vagy háttérzajjal vannak betáplálva. Az MI-asszisztensek nem minden dialektust, szlenget értenek meg. Esetenként, ha az adott MI-asszisztensnek van hozzáférése szlengszótárakhoz (amilyen például a Urban Dictionary), akkor ezzel talán könnyebben tud dolgozni. Ellenkező esetben a dialektussal kapcsolatban érzékelhető az, hogy az MI-asszisztens (például Siri, Alexa) nem érti felhasználója nyelvét. Erre példa lehet a skót, székely, palóc és sok más dialektus, ami alapvetően másképp használja nyelve fonetikáját. Emellett dialektusukban vannak olyan szavak, amelyek nem a szlenghez tartoznak, hanem terület- és kultúraspecifikusak. Itt érezhető az, hogy nem minden egyén tudja egyenrangúan használni ezeket a technológiákat, vagy nagyobb odafigyeléssel tudja csak használni ezeket. Így ezeket bizonyos közösségek nem tudják teljes mértékben kihasználni, vagy a nem megfelelő működés miatt helytelen kimeneteli eredményeket kapnak.

Carina Prunkl tanulmányában kiemeli: az egyéni autonómia külső dimenziója feltételezi a szabadságfeltételeket mint vallási és politikai szabadságot, valamint a lehetőségfeltételeket, amely a lehetőségekhez való megfelelő hozzáférést jelenti.⁴⁴ Ezen a téren előtérbe kerülnek azok az esetek is, amikor egy adott társadalmi csoport hozzáférési lehetőségei korlátozottak. A kialakulóban levő új digitális társadalom és az MI gyors fejlődésével több társadalmi csoport nem tud lépést tartani. Az ilyen helyzetben levő közösségek olyan kihívásokkal állnak szemben, mint infrastruktúra, készségek és tudáshiány, kutatási kapacitások és a helyi adatok elérhetőségének a hiánya. Itt az emberi autonómia újból érintett, mert ezt olyan külső faktorok is befolyásolhatják, mint azok a jogi, gazdasági és politikai helyzetek, melyek meghatározzák az adott közösség és társadalom életterét.

Felhasználói szabadságlimitálás

Egy másik MI-kockázat az elfogult képzés vagy az előítélet, ami a generatív MI esetében is jelen van. A képzés során ezek az adatok torzítják a válaszokat. Az NLP esetében minél sokszínűbb egy szoftver, annál jelentősebb ez a kockázat, például a kormányzati szolgáltatások, egészségügy és a HR-es interakciók esetében.⁴⁵ Az MI HR-szoftverek és MI-chatbotok jelentős hatással vannak a meglevő társadalmi konfliktusokra is, mivel a beletáplált vagy a gépi tanulás által szerzett tartalmak és adatbázisok tükrözhetik a meglevő társadalmi előítéleteket, és súlyosbíthatják is ezeket. A World Commission on the Ethics of Scientific Knowledge and Technology (COMEST) egyik 2019-es gyűlési jegyzékében kiemelte ezt a problé-

⁴⁴ Prunkl: i. m. 11–12.

⁴⁵ Stryker – Holdsworth: i. m.

mát az Amazon által használt MI-eszközzel kapcsolatosan. A munkafelvétel során ez az MI a férfi jelentkezőket részesítette előnyben.⁴⁶ Ezeket az egyenlőtlenségeket adatokból tanulja a gép, ezért a fejlesztőknek különös figyelmet kell fordítaniuk a társadalmi konfliktusok és nézeteltérések érzékeny adataira. Ebben az esetben újból előtérbe kerül a Carina Prunkl által megfogalmazott külső dimenzió,⁴⁷ mely hatással van az egyéni autonómiára. Ebben az esetben a szabadságfeltétellel találkozunk, amely az egyének és közösségek vallási és politikai szabadságát jelenti.

3.3. Az MI-alapú döntéshozatal etikai kockázatai

Az MI-alapú döntéshozatal azokat a rendszereket foglalja magába, amelyek nagy adathalmazok elemzésekor új felismeréseket és autonóm döntéseket generálnak emberi beavatkozás nélkül. Ezek a rendszerek a gépi tanulás mellett neurális hálózatokat is használnak. A gépi tanulás segít a minták azonosításában, és előrejelzéseket készít múltbeli adatok alapján. A neurális hálózatok pedig az emberi agyat utánozva rétegesen dolgozzák fel az információt, így ez lehetővé teszi a komplexebb döntések kimenetét. Ezt az MI-rendszert több ágazatban is használják, hogy növelje a pontosságot és a hatékonyságot. Így például az egészségügyi diagnosztika, a pénzügyi kérdéskör (ahol javítja a befektetési döntéseket, és felismerheti a csalásokat), a kereskedelmi előrejelzés, a személyre szabott marketing vagy az oktatás területe, ahol személyre szabja a tanulási folyamatot stb. Viszont, mint mindegyik MI-rendszer esetében, itt is találkozunk kockázatokkal.⁴⁸

A döntési folyamat átláthatósága

Az MI rendszerekbe betáplált tanulási adatok esetében láthattuk, hogy elfogultak vagy előítéletesek lehetnek. Ez a döntéshozói MI-rendszerekben is kockázatot jelent, mivel ezek az előítéletek a döntési eredményekben is megjelennek. A döntéshozatalokra kifejlesztett MI-rendszerek az átláthatóság és megmagyarázhatóság szempontjából társadalmi és felhasználói bizalmi kérdéseket is felvetnek. A felhasználók csak akkor minősíthetik át az MI által hozott döntéseket, ha átlátják a döntéshozatal algoritmusát.⁴⁹ Ez különös figyelmet igényel úgy az erkölcsi

⁴⁶ UNESCO – COMEST: *Preliminary Study on the Ethics of Artificial Intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000367823> (utolsó megtekintés dátuma: 2025. május 29.) 21.

⁴⁷ Prunkl: i. m. 11.

⁴⁸ Rinf.tech: *Ethical Considerations in AI Decision-Making*. https://www.rinf.tech/ethical-considerations-in-ai-decision-making/#elementor-toc__heading-anchor-0 (utolsó megtekintés dátuma: 2025. május 29.).

⁴⁹ Uo.

kérdések kapcsán, mint az egészségügyi és pénzügyi kérdések területén. Amennyiben az MI döntéshozatali menete nem átlátható, a felhasználója nem tudja autonóm módon eldönteni, hogy az eredmény vagy az adott döntés ténylegesen megfelel-e elvárásainak és céljainak, vagy éppen megfelel-e egyedi értékeinek. Számos MI-modell esetén, különös tekintettel a mélytanulási rendszerek jellegére, közismert a *black box*, avagy a fekete doboz fogalma, ami kevés betekintést nyújt az adott MI szoftver döntéshozatali folyamatába.⁵⁰ Az Európa Unió MI-vel kapcsolatos nyilatkozata kijelentette, hogy minden MI-rendszert úgy kell megtervezni és fejleszteni, hogy működésük kellően átlátható legyen, hogy ennek kimenetét kellően tudják értelmezni a megfelelő alkalmazás érdekében.⁵¹

Az erkölcsi döntéshozatal algoritmusra való átruházása kockázatos, amikor az algoritmus döntésének a folyamata nem átlátható és nem elemezhető. Jayesh Rane és munkatársai tanulmányukban kiemelték, hogy amikor az MI például arról dönt, hogy feltételes szabadlábra helyezzenek-e egy elítéltet, vagy hogy hitelképes-e valaki, akkor gyakorlatilag az erkölcsi tekintély egy olyan formáját gyakorolja, ami nélkülözi az etikai elvek megértését.⁵²

Felelősségvállalás

A felelősség is egy összetett kérdés az MI által hozott döntésekben. Amennyiben az MI hoz egy döntést, és felhasználója ezt ellenőrzés nélkül alkalmazza, egy helytelen döntés esetében felmerül a kérdés, hogy ki vonható felelősségre az eredményért. A felelősségre vonás kérdése a döntések meghozatalában kulcsfontosságú – már említettük, hogy a felhasználó autonóm döntését befolyásolhatja egy MI által hozott döntés.⁵³ Ilyen esetekben gondolhatunk egy szakorvosra, aki karrierje során sok esettel találkozott, és pár vizsgálat után már sejtheti páciensének az egészségügyi gondját, viszont még több vizsgálatot kér, hogy döntését megalapozza. Ilyen esetben, ha az MI döntéshozatali folyamata nem visszatekinthető vagy elemezhető, a szakorvos nem tudja autonóm módon meghozni a döntését.

⁵⁰ Rane, Jayesh et al.: Enhancing black-box models: Advances in explainable artificial intelligence for ethical decision-making. *Future Research Opportunities for Artificial Intelligence in Industry 4.0 and 5.0*. 2024. (136–180.) 137.

⁵¹ Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete: A mesterséges intelligenciáról szóló rendelet, 13. cikk, 1–6. cikkely, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689> (utolsó megtekintés dátuma: 2025. május 29.).

⁵² Rane et al.: i. m. 144.

⁵³ Rinf.tech: i. m.

A felelősség kérdésköre jelen van a félautomata és önvezető autók esetében is. A számítógépes látás és döntéshozatali algoritmusoknak köszönhetően több olyan technológiával is találkozunk, amelyek a felhasználók körében elterjedtté váltak: Advanced Driver Assistance System (ADAS), avagy a fejlett vezetőtámogató rendszerek; Human–Machine Transition (HMT), avagy az ember–gép átmenet ciklus; Pay As You Drive (PAYD), avagy „úgy fizetsz, ahogy vezetsz”.

Különböző esetekben a vezető autonómiáját meghatározza az automatizálás: például a vezető képességei elhomályosodhatnak, mert túlzottan támaszkodik gépjárműve ADAS-technológiájára. A HMT-ciklus kiemelt jelentőségű, mivel az emberi vezető és a rendszer felületei közötti interakcióval foglalkozik. A vezető igénye és a vezetési helyzet kritikussága szerint az interakció több szinten működik: információ közlése, figyelmeztetés, a járművezérlés átadása vagy átvétele. Ez a folyamatos értékelés és felügyelet különböző lehetőségeket nyújt a biztosítótársaságoknak. A Pay As You Drive (PAYD), avagy „úgy fizetsz, ahogy vezetsz” logika a biztosítási díj változó, a járművezető teljesítményétől függő árazást jelenti.⁵⁴

A felelősségvállalás itt a vezetői autonómia kérdéskörében problémákat vet fel mind az ADAS, mind a HMT esetében. Nyilvánvaló, hogy amennyiben törvényteleség vagy baleset történik, a vezetőt fogják felelősségre vonni, és nem gépjárműve döntéseit. Ebben az esetben a felelősség a vezetőre hárul. A 2017-es németországi jogszabály előírta, hogy a vezetőnek mindig ébernek kell lennie, hogy „időben” átvehesse járműve felett az irányítást.⁵⁵

3.4. A vizuális analitika, számítógépes látás és mintafelismerés etikai kockázatai

A számítógépes látás, vizuális analitika és mintafelismerés az MI azon területei, amelyek egy másik szempontból alakítják a technológiai fejlesztéseket és a technológiák alkalmazását. A számítógépes látás etikai megközelítése magába foglalja a proaktív intézkedéseket, felelős gyakorlatokat és a folyamatos ellenőrzést. A számítógépes látás az MI-technikák olyan felhasználását jelenti, ahol a képek és videók elemzése történik. Így ez számos feladatot foglal magába: követés, osztályozás, lokalizálás, szegmentálás, arc-, érzelem- és viselkedésfelismerés. Ez hasz-

⁵⁴ Bellet, Thierry et al.: From semi to fully autonomous vehicles: New emerging risks and ethico-legal challenges for human-machine interactions. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2019/63. (153–164.) 158–159.

⁵⁵ Uo. 159.

nálható a kézzel írott szöveg olvasásától a közlekedési táblák felismeréséig vagy a közönség hangulatelemzéséig is.

Adatvédelem

Az egyéni autonómia tiszteletben tartása szükséges az etikus számítógépes látáshoz, vizuális analitikához és mintafelismeréshez. Ez azt jelenti, hogy a technológia és a technológiával rendelkező „erő” tiszteletben kell hogy tartsa az emberek szabadságát és jogát arra, hogy saját adataik felett rendelkezzenek, és döntsenek ezek felhasználásáról. Az arcfelismerő és képelemző technológiák alkalmazásához szükséges lenne az egyén tájékozott beleegyezése. Az egyénnek egyértelmű magyarázattal és pontos felhasználási móddal kellene szolgálni, hogy adott esetben mire fogják ezt a technológiát használni. Találkozunk olyan esetekkel, ahol az adott szoftver részben tájékoztatja felhasználóját jogairól, a kockázatokról és a felhasználási célokról, viszont a felhasználó nem minden esetben veszi a fáradságot, hogy elolvassa és értelmezze a tájékoztatót.

Rosalie A. Waelen tanulmányában beszél a diszpozíciós hatalomról, ami egy bizonyos képességet jelent. Ezzel azt emeli ki tanulmányában, hogy az egyén vagy csoport lehetőségei annak függvényében nőnek vagy csökkenek, hogy lehetőséget nyernek-e bizonyos cselekvésre, vagy pedig attól megfosztva maradnak. Így a nagyobb hatáskör elősegíti az autonómiát. A szerző értelmezésében az autonómia azt jelenti, hogy az emberek jobban és könnyebben képesek úgy cselekedni, ahogyan azt ők jónak látják, szabadságuk és lehetőségük van fejleszteni önmagukat és felfedezni lehetőségeiket.⁵⁶ Így ez kapcsolódik Carina Prunklnek ahhoz a megközelítéséhez, amely során az autonómia második dimenziójával foglalkozik, és megfogalmazza, hogy ez nem más, mint a szabadságfeltétel és a lehetőségfeltétel.⁵⁷

CCTV kamerák és felhasználásuk

Annak ellenére, hogy törekvések vannak az átláthatóságra és az adatvédelemre, az MI-t több forrásból is képezik különféle célokra. Sok esetben az ilyen képzési anyagokat a nyilvános közösségi tereken levő CCTV kamerák segítségével gyűjtik be. Ha az egyén tudna is róluk, nehezen kerülhetné el őket, mivel ez azt jelentené, hogy sok közösségi téren nem mozoghatna nyilvánosan, nem használhatna bizonyos telefonos alkalmazásokat stb. Maga az információs kontroll és a

⁵⁶ Waelen, A. Rosalie: The ethics of computer vision: An overview in terms of power. *AI and Ethics*, 2024/4. 353–362. 355.

⁵⁷ Prunkl: i. m. 11–12.

tudatlanság az emberi autonómia ellen szól, mivel korlátozza képességét és az ítélet alkotásához szükséges információit.

Ezenkívül a számítógépek illetéktelen látása jogi szempontból is kifogásolható, hisz ez korlátozza az emberi önrendelkezést, mivel az egyén nem mer autonóm módon cselekedni. Maga a szabadság és a lehetőségek meghatározhatják az egyén autonóm döntéseit és cselekvéseit. Carina Prunkl tanulmányában kiemeli, hogy az arcfelismerő technológia egy hatékonyabb megfigyelést tesz lehetővé, ahol az MI csak segítő szerepet tölt be. Ez a technológia az emberek szabadságjogának az aláásására is használható, mivel itt az egyén autonómiáját érinti a tömeges megfigyelés.⁵⁸ Egy példa az emberi autonómia mint szabadságjog megsértésére a Kínai Népköztársaság megfigyelési rendszere. Pekingben például arcfelismerő technológiákat szereltek fel több közösségi térre. Emiatt sok kínai állampolgárnak aggályai vannak, mivel úgy gondolják, hogy ezek a rendszerek nem nyújtanak megfelelő adatvédelmet, és nyomon követik az emberek mozgását. Másfelől az arcfelismerő közösségi média kritikái is árulkodóak, miszerint nem veszélytelen az, hogy „a Nagy Testvér szemmel tart téged”.^{59,60} A Kínai Népköztársaság arcfelismerő rendszere nemcsak nyomon követi a polgárokat, hanem próbálja befolyásolni viselkedésüket is. Erre a gyakorlatra egy példa, amikor a kínai arcfelismerő rendszer automatikusan felteszi egy polgár nevét és fényképét egy hirdetőtáblára, ha szabálytalanul közlekedett, privát üzenetben pedig bírságot is kap.⁶¹ A szabálytalan zebraátkeléstől és ennek publikus megosztásától eltekintve is az állampolgárok tisztában vannak azzal, hogy megfigyelik és büntetik őket, ha nem az adott társadalom elvárásai és törvényei szerint cselekednek. Itt a félelem és kötelességtudat miatt az egyén autonóm cselekedete és erkölcsi döntése nem a sajátja lesz, mert kényszerből egy követési, megfelelési mechanizmushoz alkalmazkodik.

Rosalie A. Waelen kiemeli, hogy a számítógépes látás esetében, legyen az videokamera, utcai vagy telefonos kamera, a túlzott felügyelet és megfigyelés is hatással van az autonómiára, pontosabban az erkölcsi ítélőképességre és cselekedetre.

⁵⁸ Uo.

⁵⁹ “Big Brother is keeping an eye on you” – idézet George Orwell 1984 című regényéből. A regény a totalitárius diktatúrarendszerek megfigyelési módjának negatív utópiája.

⁶⁰ Dauvergne, Peter: The everyday politics of facial recognition in China. In: Dauvergne, Peter (ed.): *Political Science and Public Policy*. Edward Elgar Publishing Ltd., Northampton, 2022. (69–75.) 70–72.

⁶¹ NG, Alfred: How China uses facial recognition to control human behavior. <https://www.cnet.com/news/politics/in-china-facial-recognition-public-shaming-and-control-go-hand-in-hand/> (utolsó megtekintés dátuma: 2025. május 29.).

A számítógépes látás lehetővé teszi, hogy a legújabb személygépkocsiban levő kamerás megfigyelés megítélje a személy vezetési stílusát, amivel a biztosítótársaság célja, hogy ösztönözze a sofőröket a biztonságos vezetésre. A munkáltatók számára is lehetőség nyílik alkalmazottjaik megfigyelésére, így a munkatársakkal való csevegést és az íróasztalnál töltött időt is számontartják.⁶² Hasonló megfigyelésre már volt példa az MI-t megelőzőleg is, ahol az otthon dolgozó alkalmazottakat aszerint ellenőrizték, hogy mozgott-e az egér vagy sem. Itt a túlzott ellenőrzés újból az ember autonómiáját érinti, mivel ez a követési rendszer nem az egyén sajátos cselekedeteit hozza felszínre, hanem azokat, amelyeket automatikusan a megfelelés miatt tesz.

A számítógépes látás hat a személyes autonómiára azzal, hogy befolyásolja vagy meghatározza az ember lehetőségeit, azt, hogy milyen értékeket fogad el, és milyen értékek szerint él. Ezenfelül a számítógépes látást alkalmazó technológiák visszatükrözik vagy súlyosbítják a különböző társadalmi problémákat, például a faj, nem, bőrszín stb. miatti előítéleteket. Az előítéletes algoritmus problémáját már tárgyaltuk az előző fejezetben, viszont ebben az esetben is visszaköszön, hiszen a számítógépes látás hamarabb és nagyobb valószínűséggel ismeri fel a fehér embereket és a férfiakat, mint a sötétebb bőrűeket és a nőket.⁶³ Ez az algoritmus az általa negatívan érintett emberek önfejlesztését és önértékelését csökkenti és degradálja.

Mindemellett egy számítógépes látás modell – pontosabban az arcfelismerő és elemző szoftverek – célja az emberek azonosítása és kategorizálása. Az ilyen szoftvereket számos célra lehet használni: gyanús egyének felismerésére, vásárlói réteg beazonosítására, vicces arcfilterek létrehozására vagy éppen tömeges megfigyelésre.

Ügyszintén egy másik gyakorlat a Kínai Népköztársaságban – amely a tömeges megfigyelés eredménye az arcfelismerő technológiáknak köszönhetően – egyes kínai társadalmi kisebbség üldözése vagy éppen „népirtása”. A népirtás célja, hogy egy csoportot etnikumként, vallásként és/vagy kultúraként támadjon, ami példánk esetében az ujugrokat és más muszlim csoportokat jelenti. Esetünk folyamata viszont jelentősen különbözik más népirtásoktól. Több stratégiát alkalmaznak, amelyek közül az egyik, hogy Kína gyáraiban dolgoznak minimális, jelképes összegért, elszakítva őket gyermekeiktől és otthonuktól.⁶⁴ Ahogy az arcfelismerés technológiája fejlődött, több kamera technológiájának a tesztelése is szerepet játszott az ujgur és más etnikai kisebbségek nyomon követésében és esetleges őrizetbe vételében.

⁶² Waelen, i. m. 357.

⁶³ Uo. 359.

⁶⁴ Turdush, Rukiye – Fiskesjö, Magnus: Dossier: Uyghur Women in China's Genocide. *Genocide Studies and Prevention: An International Journal*. Vol. 15/1. 22–43/2021. 31–32.

ben. Több ilyen kameragyártó cég terméke már 2019-ben is képes volt az ujjurok és nem ujjurok osztályozására.⁶⁵

A szegény szociáletikai és társadalometikai aspektusait vizsgálva szeretném kiemelni az egyéni autonómiára gyakorolt hatásokat. Fazakas Sándor *Új erkölcsi imperatívusz: Ne szegyenkezz... és senkit meg ne szegyeníts! A szegény teológiai értelmezése és szociáletikai relevanciája* c. tanulmányában⁶⁶ kiemeli a szegény szociáletikai aspektusait, melyek relevánsak az egyéni autonómiára gyakorolt hatások terén. Az említett aspektusokból célom kiemelni a megszegyenítést mint hatalmi eszközt.

A fenomenológiai, pszichoanalitikai és teológiai értelmezés esetében kiemelte, hogy az az egyén, akit valamilyen módon megszegyenítenek, szeretne elrejtőzni, eltűnni, majd végső esetben önmagát is feladni. Ez a külső értékítélet romboló hatást gyakorol az egyén önértékelésére, önértetére, végül a külső értékítélettel való teljes vagy részleges azonosuláshoz vezet.⁶⁷ Így a megszegyenítő valamilyen módon, akár verbálisan, akár manipuláció révén korlátozza vagy negatív irányba változtatja meg az egyén önmagáról kialakított képét, és így – ha az illető azonosul ezzel a külső véleménnyel – énképe sérül. Ilyen körülmények között, úgy véli, „nem lehet az, aki vagyok, mert megfosztanak attól, ami a személyiségem belső tartalmát képezi, s pellengérre kerül az, aminek védve-rejtve kellene maradnia”.⁶⁸ Ezúton a megszegyenítőnek hatalma lesz az „áldozat” felett, mivel sérti integritását, autonómiáját. Ha visszatérünk a zsidó hagyományokhoz és a Talmudhoz, a következő intelmmel találkozunk: „Aki a másik embert nyilvánosan megszegyeníti, olyan, mintha vért ontott volna.”⁶⁹ Ezt az intelmet Fazakas Sándor úgy hangsúlyozza, mint új erkölcsi imperatívuszt, és sürgeti, hogy a „másik ember megszegyenítéstől való megóvása a társadalmi élet elementáris kötelezettsége legyen”.⁷⁰ Máskülönben az egyént megfosztják saját értékétől, autonómiájától, meggyőződéseitől. Ez az MI-technológia által digitalizált környezetben pedig újból hangsúlyos szerepet kap, mivel az „áldozat” megszegyenítésének tanúja lehet a virtuális tér számtalan felhasználója.

⁶⁵ Bhuiyan, Johana: 'There's cameras everywhere': Testimonies detail far-reaching surveillance of Uyghurs in China. <https://www.theguardian.com/world/2021/sep/30/uyghur-tribunal-testimony-surveillance-china> (utolsó megtekintés dátuma: 2025. május 29.).

⁶⁶ Lásd. Fazakas Sándor: *Új erkölcsi imperatívusz: Ne szegyenkezz... és senkit meg ne szegyeníts! A szegény teológiai értelmezése és szociáletikai relevanciája*. *Studia Universitatis Babeş-Bolyai, Theologia Reformata Transylvanica*, 2018/63/1. 19–37.

⁶⁷ Uo. 29.

⁶⁸ Uo.

⁶⁹ Goldschmidt, Lazarus: *Der babylonische Talmud*. Jüdischer Verlag, Berlin, 1993, 634.

⁷⁰ Fazakas: *Új erkölcsi imperatívusz*. 29.

4. Összegzés és kitekintés

Következtetésképpen látható, hogy az egyéni autonómia az MI korszakában olyan komplex feladat, amely interdiszciplináris megközelítést igényel. Az autonómia fogalomkörének a tárgyalásában említettem az autonómia és moralitás kapcsolatát, valamint kiemeltém Németh Dávid azon érvét, hogy a társadalmi kapcsolatok működőképességét egyedül a jogrendszer nem képes fenntartani, hiszen az igazi összetartó erőt a moralitás adja.⁷¹ Ebből a megközelítésből egyértelművé válik, hogy az erkölcsi nevelésnek kulcsfontosságú szerepet kell szánni a társadalomban és a családban. Az oktatás és nevelés esetében a nevelési folyamatok egyben heteronóm folyamatok is, mivel igyekszik nevelni, oktatni és irányítani az egyén személyiségének a fejlődését. Másik irányból pedig fontos kiemelni, hogy azért törekszik erre, hogy az egyén személyiségéből fakadóan tudjon önálló erkölcsi döntéseket hozni.⁷²

A teológiai etika alapjaira visszatérve érdemes hangsúlyozni Szűcs Ferenc értelmezését, aki – Barth kifejezésével – „speciális” etikáról beszél, amely azt vizsgálja, hogy az ember hogyan válaszol Isten tetteire engedelmisséggel vagy engedetlenséggel. Így hangsúlyozza, hogy a teológiai etika sajátossága, hogy az az „isteni cselekvés által meghatározott emberi cselekvés tudománya”.⁷³ Nyilvánvaló, hogy a társadalom és az egyének folyamatosan kihívásokkal és kockázatokkal szembesülnek, és ezekre megoldásokat kell találniuk, hogy felelősségteljes, erkölcsileg helytálló döntéseket hozhassanak. Ezek a döntések viszont nem minden esetben hozhatók meg pusztán gyakorlatias megfontolások alapján.⁷⁴ Mindemelllett a mai társadalom individualizálódott, így az erkölcsi szféra is valamiképpen privatizálódott, ám az együttélés mégsem korlátozódhat csupán a jogi szabályozásra, szükség van a moralitásra is – ahogyan ezt okkal hangsúlyozza Németh Dávid.⁷⁵ Szűcs Ferenc viszont kiemeli, hogy a legrosszabb, amit a teológiai etika tehet, az a moralizálás és az apokaliptikus fenyegetőzés eszköztárának a használata. Ehelyett feladata felmutatni az „egyéni élet páratlan, Isten szerinti értékét, hangsúlyozva azt is, hogy az emberlét alapformája a személyes Én–Te kapcsolatban rejlik”.⁷⁶ Ennek hármas vonatkozása van: az Istennel, önmagunkkal és a felebarátunkkal való kapcsolat.

⁷¹ Németh: i. m. 27–29.

⁷² Szűcs, i. m. 26.

⁷³ Uo. 11.

⁷⁴ Uo. 16.

⁷⁵ Németh: i. m. 27–29.

⁷⁶ Szűcs: i. m. 19.

A teológiai etika célja felmutatni a keresztyén szabadság páratlan lehetőségeit, amelyek egyszersmind felelősséggel is járnak.⁷⁷

Láthattuk, hogy az etika mint erkölccsel foglalkozó tudomány már az ókortól vizsgálta, hogy milyen cselekedetek tekinthetők erkölcsösnek, igazságosnak, és azt, hogy mi alapján ítéljük meg a jót és a rosszat. Elemzésünkben előtérbe került a Szűcs Ferenc által tárgyalt hármas rendszer: indíték, cél és kiváltott hatás.⁷⁸ Az MI-technológia korában az egyének olyan kihívásokkal és kockázatokkal szembesülnek, melyekre megoldásokat kell találniuk, és felelősségteljes döntéseket kell hozniuk. Ennek értelmében elengedhetlenné válik a racionális érvelés és kritikus gondolkodás, az erkölcsi meggyőződés kialakítása és az erkölcsi ítéletalkotás, melynek egyik indítéka az intuíció. Az intuíció esetében szükséges kiemelni az azokat a meghatározó elemeket: az emlékek, élmények, tapasztalatok sokaságát és egyéb, a szocializáció során kialakított, áthagyományozott vagy elsajátított cselekvési minták sorát. Ehhez úgyszintén hozzátartoznak az egyén viselkedésmintái vagy a neveltetése során kialakított erkölcsi meggyőződései is.⁷⁹

A téma komplexitásának köszönhetően ezzel egyidejűleg előtérbe került a cselekedetek szándékossága és indíttatása, valamint az autonómia és heteronómia kérdésköre. Ezért figyelembe kell venni az autonómia belső és külső dimenzióit. Ebben az esetben a belső dimenzió nem más, mint az egyén belső meggyőződéseinek, értékeinek, motivációinak és okainak a szférája, mely feltételezi, hogy ez csakis az egyén tulajdona és nem valamilyen külső hatás terméke. A külső dimenzió pedig kiemeli az egyén szabadság- és lehetőségfeltételét.⁸⁰ Amennyiben a külső dimenzió nem teljesül, úgy az egyén nem tudja autonómiáját kamatoztatni. Az MI által támogatott technológiák esetében láthatóvá vált, hogy az egyén autonómiáját milyen kockázatok veszélyeztetik, akár belső, akár külső szinten.

Több MI-technológia esetében találkozunk azzal, hogy külső dimenzióként hatnak az egyének autonómiájára a lehetőség- vagy szabadságfeltétel limitálása-ként. Viszont szükséges kiemelni, hogy a felhasználók a szabadság- és lehetőségfeltételeknek köszönhetően alkalmazzák ezeket a technológiákat, így azok direkt vagy indirekt módon hatnak rájuk. Találkozunk olyan szabadság- és lehetőségfeltétel-sértéssel, amikor a megszegényítést hatalmi eszközként kezelik. Mind az embertársi, mind a hatalmi eszközként használt cselekedetek révén az egyéni autonó-

⁷⁷ Uo.

⁷⁸ Uo. 2.

⁷⁹ Fazakas: *A protestáns etika kézikönyve*. 46–47.

⁸⁰ Prunkl: i. m. 11.

mia sérül, mivel a megszégyenítő egyén vagy entitás valamilyen módon korlátozza és negatív irányba változtatja az egyén önmagáról kialakított képét, így sérül annak identitása. Ily módon akarata ellenére megváltozik megítélése, mivel gyengeségeit és hiányosságait ellene fordítják.

A másik oldalon találkozunk azzal a jelenséggel, amikor belső dimenzióként lehet megfigyelni, hogy miként hatnak az MI-technológiák az egyén autonómiájára. Ezek lehetnek technológiai, fejlesztői, felhasználói kockázatok egyaránt. Mindamellett, hogy az erkölcsi nevelés és a moralitás kulcsfontosságú szerepet kap, fontos kiemelni az egyének azon racionális és kritikus gondolkodási képességeit, amelyekkel a kockázatokat limitálhatják önmagukra nézve, valamint amelyek révén képesek a cselekedeteiket megelőző erkölcsi meggyőződésre és ítéletalkotásra. Ezúton előtérbe kerül az intuíció is mint az erkölcsi döntések olyan indítéka, amely számára meghatározóak az egyén élettapasztalatai és a szocializációja során elsajátított vagy áthagyományozódott értékei és cselekvési mintái. Így előtérbe kerül a társadalom és a közösség szerepe az egyén fejlődése során mint meghatározó tényező, valamint az erkölcsi nevelés fontossága is. Okkal érvel Németh Dávid⁸¹ az erkölcsi nevelés fontossága mellett, mivel a mai társadalomban az általános moralitás fokozatosan romlik, és ezt a szerepet a családok és közösségek egyre kevésbe tudják betölteni.

Végül az egyéni autonómiára ható külső dimenziók szintjén fontos szerepet nyer a Fazakas Sándor által megfogalmazott új erkölcsi imperatívusz, miszerint „Ne szégyenkezz... és senkit meg ne szégyeníts!”⁸² – ezzel is óvhatjuk embertársunk identitását, integritását, értékeit és énképét a sérülésektől. Az is egyértelmű, hogy felértékelődik az erkölcsi nevelés fontossága, mely döntő tényező az egyén személyiségének a fejlődésében, mivel ezáltal lesz képes az önálló, erkölcsi döntések meghozatalára. Az ehhez kapcsolódó racionalitás és kognitív képességek elősegítik azt, hogy az egyén az erkölcsi meggyőződés útján megoldásokat találjon és felelősségteljes döntéseket hozzon. Az erkölcsi nevelés és a közösség hatása tehát kiemelkedő szerepet kap az egyén erkölcsi fejlődésében, ezzel együtt pedig a felelősségteljes döntések rejtett indítékának, az intuíciónak a formálódásában is.

⁸¹ Németh: i. m. 29.

⁸² Lásd Fazakas: *Új erkölcsi imperatívusz*. 19–37.