

Pécsi Tudományegyetem
Filozófia Doktori Iskola

DISSZERTÁCIÓ

Szilágyi Tamás

2024

PÉCSI TUDOMÁNYEGYETEM
FILOZÓFIA DOKTORI ISKOLA

SZILÁGYI TAMÁS

**FELELŐSSÉG ÉS INNOVÁCIÓ
A TUDOMÁNYOS FEJLŐDÉS ETIKAI ASPEKTUSAINAK
VIZSGÁLATA**

Disszertáció

Témavezető:
prof. Dr. Nemeskéri Zsolt
egyetemi tanár

Pécs

2024

Tartalomjegyzék

I. Bevezetés	3
1.1 Módszertani megjegyzések	9
II. Fogalmi konceptualizáció	11
2.1 Az alkalmazott filozófia, mint teoretikus keret	11
2.2 Az alkalmazott etika sajátosságai	16
2.3 Felelősség a nyugati gondolkodásban	22
2.3.1 Szabadság, ágencia és felelősség	22
2.3.2 Az erkölcsi felelősség fogalmának társadalmi és nyelvi vonatkozásai	25
2.3.3 Richard McKeon a felelőségről	26
2.4 Az innováció értelmezésének lehetőségei	30
2.5 Összefoglalás	39
III. Az innováció filozófiatörténeti kontextusa	41
3.1 Az újdonsághoz való viszony az ókori görög filozófiában	41
3.1.1 Platón viszonya az újításokhoz	43
3.1.2 Arisztotelész Hippodamosz államelmélet kritikája	44
3.2 Az újítás szerepe a középkori arab filozófiában	48
3.3 Aquinói Szent Tamás a törvények megváltoztatásáról	52
3.4 Az innováció, mint az adaptáció eszköze Machiavelli államelméletében	53
3.5 Az új dolgok hasznosságáról: Francis Bacon az innovációról	55
3.6 A társadalom változásorientáltsága Giambattista Vico filozófiájában	59
3.7 Marx és a technológiai forradalom dialektikája: az innováció ígérete és paradoxonjai	61
3.8 Nietzsche hatása a teremtő rombolás elméletére	64
3.9 Összefoglalás	65
IV. Az innováció az erkölcs és a technológiai fejlődés ütközőzónájában	68
4.1 A tudomány új horizontja: az innovációvezérelt kutatás lehetőségei és kihívásai	68
4.2 A felelősség és az ágencia kérdései az innovációetikában	69
4.3 A felelős kutatás és innováció etikai keretei: közösségi és egyéni perspektívák	78
4.4 Az etika, az innováció és a felelősség összefüggései a kortárs tudományosságban	81
4.5 A felelőségetika mint az innovációetika normatív keretrendszere	93
4.6 Az innovációs folyamatok felelőségetikai irányelvei	99
V. A mesterséges intelligencia etikai és társadalmi kihívásai	101
5.1 Felelős innováció és etikai megfontolások az AI-fejlesztésben	101
5.1.1 A mesterséges intelligencia etikai szabályozásának fejlődése	102
5.1.2 Nemzetközi irányelvek és szabályozási törekvések az AI területén	103
5.1.3 Összehasonlító elemzés: AI etikai irányelvek és szabályozások	109

5.1.4 Az etikai szabályozás hiányosságai.....	114
5.2 Etikai kihívások a mesterséges intelligencia alkalmazásának területén.....	117
5.2.1 Az innováció kihívásai a biztonságpolitika területén.....	117
5.2.2 A globális adathálózatok etikai kihívásai.....	123
5.2.3 A mesterséges intelligencia és az automatizáció hatása a munka világára	129
5.2.4 A mesterséges intelligencia alkalmazásának etikai vonatkozásai az oktatásban ..	132
VI. A felelős döntéshozatali mátrix, mint az etikai értékelés keretrendszere	137
6.1 A félelem konstruktív becsatornázása.....	137
6.2 A félelem heurisztikája, mint a felelős döntéshozatali mátrix alapelve.....	139
6.2.1 Értékelési rendszer	141
6.2.2 Adaptációs esettanulmány a felelős döntéshozatali mátrixhoz	149
6.3 Összefoglalás.....	155
VII. Összegzés és konklúzió	156
Köszönetnyilvánítás	161
Felhasznált irodalom	162
Internetes forrás.....	173
Ábrák jegyzéke.....	175
Táblázatok jegyzéke.....	175
Mellékletek.....	176

I. Bevezetés

Az amerikai filozófus Carl Mitcham, akinek a munkássága jelentős mértékben hozzájárult a technológia filozófiai és etikai vonatkozásainak megértéséhez, a *Thinking Through Technology: The Path between Engineering and Philosophy* (Mitcham, 1994) című nagyhatású könyvének egyik központi gondolata az, hogy a technológia nem csak funkcionális eszközök összessége, hanem az emberi társadalmakat és gondolkodást alapjaiban befolyásoló és meghatározó komplex kulturális jelenség, aminek megértéséhez egy olyan kritikai megközelítésre van szükség, amely etikai dimenziókat is magában foglal.

Az első és a második ipari forradalom idején a gőzgépek, az elektromosság és a tömegtermelés alapjaiban változtatták meg a mindennapi életet és a hadviselést, míg a 20. század első évtizedeiben a repülőgép, a rádió, a penicillin megjelenése hozott hasonlóan radikális átalakulást. A technológiai fejlődés napjainkban tapasztalható felgyorsulása nem előzmények nélküli, sok hasonlóságot mutat a korábbi innovációs korszakokkal, ugyanakkor fontos különbségek is kirajzolódnak. Ezek közül az egyik legfontosabb, hogy a jelenlegi hullám üteme sokkal gyorsabb és kiszámíthatatlanabb, míg a korábbi korszakokban sokszor évtizedek teltek el két nagy áttörés között, ma viszont számos innováció követi egymást villámgyorsan, összetett rendszereket alkotva. Fontos kiemelni továbbá, hogy az összekapcsoltság révén, az egyes fejlesztések ma a globális környezetben méretik meg magukat és terjedésük is hamar túllép a regionális vagy nemzetállami kereteken. A modern technológiai innovációk – különösen a digitális kommunikációs eszközök megjelenése és a közlekedési eszközök fejlődése – radikálisan megváltoztatták az emberek, az információk és áruk mozgását, ami lehetővé tette a nemzeti határok jelentőségének csökkenését és egy globálisabb társadalom kialakulását. A nemzetállam, mint a jelent meghatározó, domináns politikai és társadalmi szerveződési forma viszonylag új jelenség a történelemben. Korábbi időszakokban, mint a Római Birodalom, illetve később az Oszmán vagy a Brit Birodalom idején, a technológiai és kulturális transzfer eredendően határokon átívelő és nemzetközi volt. Vagyis ebben az értelemben a modern globalizáció és a nemzeti határok jelentőségének csökkenése kapcsán, lényegében egy korábban már létező és civilizációs köröktől függetlenül széles körben uralkodó helyzethez térünk vissza, egy integráltabb és összekapcsoltabb világ felé mozdulva el.

Mindez döntéshozói szinteken nagy rugalmasságot követel, mivel egyszerre válik szükségessé mérlegelni egy cselekvés várható gazdasági, társadalmi és politikai következményeit, tudománypolitikai megfontolásait, kulturális szenzitivitását és az uralkodó morális keretrendszerek normáihoz való viszonyát.

A technológia *apokaliptikus potenciáljára* (Jonas, 1984, p.45) tekintettel, az innovációs tevékenységek generálta kockázatok problémája kiemelt kérdés lett napjainkban. A technológiai fejlődés újradefiniálja a biztonság hagyományos paradigmáit, potenciális fenyegetések és lehetőségek új dimenzióit nyitva meg. Az innováció és a politika területének növekvő konvergenciája miatt, a nemzetállamok csakúgy, mint technológiai fejlesztésekben érintett piaci szereplők, kettős kihívással küzdenek: egyfelől, hogy a legszélesebb területeken kiaknázzák az újítások előnyeit, valamint ezzel párhuzamosan megőrizték biztonsági pozícióikat, rugalmasak és reziliensek maradjanak a szinte napról-napra jelentkező kihívásokkal szemben (Wanzenböck et al., 2020). A szűkebb értelemben vett technológiai innováció gyakran olyan fejlesztéseket foglal magában, amelyek egyszerre erősítik a biztonsági rendszereket és generálnak új kihívásokat – etikai területeken is. A technológiai innovációk elterjedésével – a mesterséges intelligenciától a blokkláncig – a globális környezet radikális átalakuláson megy keresztül.

Ebben a globális kontextusban az emberi felelősség egyre nagyobb figyelmet kap, a tudományos tevékenységek szabályozásának és szabályozhatóságának kérdésköre kiemelt jelentőségű témává válik a kortárs etikai diskurzusban. Az új technológiák – mint a mesterséges intelligencia – terjedése és például az egyre növekvő kiberbiztonsági kockázatok, jelentősen befolyásolják az emberi élet számos aspektusát, ezért elengedhetetlen, hogy a döntéshozók figyelembe vegyék ezek társadalmi hatásait, és a szabályozás etikai alapú lehetőségeit. Az etikai megfontolások nem pusztán a jogi keretek kialakítását segítik elő, hanem hozzájárulnak a bizalom és az együttműködés megerősítéséhez is a globális közösségben.

Az innováció kérdésköre eredendően nem tárgya a filozófiának – ahogy ezt a harmadik fejezetben részletesen tárgyalom majd –, azonban az innováció ütemének felgyorsulása és hatókörének kiszélesedése olyan változásokat idéz elő a tudományban, valamint közvetlen módon a társadalmakban, amelyek már konkrét etikai, episztemológiai és ontológiai kérdéseket vetnek fel. Ki viseli a felelősséget egy innováció nem szándékolt következményeiért? Hogyan osztható meg a felelősség az innovációs folyamat különböző szereplői között? Mennyire vagyunk képesek előre látni, elemezni és kiértékelni egy innováció hosszú távú következményeit? Milyen fogalmi bázison leszünk képesek az új jelenségek leírására, megértésére és interpretálására? Milyen ontológiai státusszal rendelkeznek a mesterséges intelligencia által létrehozott entitások? Miként alakítja át az innováció a valóság természetéről alkotott felfogásunkat? Ezek a kérdések túlmutatnak a pusztán technológiai vagy gazdasági megfontolásokon és diszciplináris kereteken, és vitathatatlanul filozófiai természetűek.

A filozófiai vizsgálat létjogosultságát megerősíti továbbá, hogy az innovációs folyamatok nem légüres térben zajlanak, hanem beágyazódva összetett társadalmi, gazdasági és ökológiai rendszerekbe – ezek kölcsönhatásainak megértésében a filozófia egy holisztikus perspektívát és széleskörű módszertant kínál. Az alkalmazott filozófia és etika lehetővé teszi, hogy valós problémákra alkalmazzuk a filozófiai gondolkodás eszköztárát, keretet biztosítva a kritikai reflexióhoz és az erkölcsileg megalapozott döntéshozatalhoz. Nyíri Kristóf a *The Humanities in the Age of Post-Literacy* címet viselő 1996-os tanulmányában a kommunikáció felől közelítve leírja, hogy a bölcsész tudományok és a filozófia jövője összefonódik az új médiumok előretörésével és a hálózatosodással, a jelentés helyett mindinkább az alkalmazásra áttevődő fókusszal. Ebben a kontextusban a filozófia új szerepe abban áll, hogy segít értelmezni ezt az átalakuló tudástérképet, kritikusan vizsgálva az új technológiák (és kommunikációs formák) hatását gondolkodásunkra és társadalmunkra, miközben ötvözi a hagyományos filozófiai kérdésfelvetéseket az új médiumok által felvetett kihívásokkal (Nyíri, 1996, pp. 110-116).

Az innováció filozófiai szempontból történő vizsgálata ma nem pusztán tudományos feladat, egyúttal törekvés a társadalmi-politikai folyamatok *befolyásolására* is. A filozófia *alkalmazója* nem értéksemleges szemlélő, hanem *normaalkotó* és *véleményformáló* is egyúttal, a közös jövő alakítója, ahogy Lányi megjegyzi:

„A jelenben zajló társadalmi és természeti folyamatok trendjeit a velük kapcsolatos kulturális attitűdök és empirikusan vizsgálható jövőképek összefüggésében tanulmányozza, annak tudatában, hogy az értelmezés keretei (és eredménye) egyáltalán nem függetlenek saját, a jövőre irányuló intencióitól. Azaz jövőképe nem „kép” a jövőről a szó szorosabb értelmében, hanem a jövőre irányuló cselekvési program része. A jövőképkészítés akarva-akaratlanul befolyásolja az elkövetkezendőket, s jobb, ha ennek világos tudatában történik, mint ha a tudományos vizsgálódás semlegességére hivatkozva elhárítani igyekszik ebből adódó felelősségét” (Lányi, 2007, p. 8).

Az innováció mibenlétéről, természetéről és a szabályozásának lehetőségeiről szóló diskurzus kiszélesítése a kortárs filozófia számára többé nem periférikus, hanem mindinkább központi kérdés. A gyors technológiai fejlődés új kihívásokat és lehetőségeket teremt, amelyek megkövetelik az etikai normák újragondolását és új szemléletmódok adaptálását, különösen a tudományos kutatás és fejlesztés egyre nagyobb mértékben specializálódó területén.

A döntés, hogy egy filozófiai disszertációban az innovációra összpontosítsak, nem véletlenszerű tehát, hanem annak a szükségletnek a tükröződése, hogy a gyors és beláthatatlan

következményekkel járó változások korában kritikai reflexióra és etikai útmutatásra van szükség.

A doktori disszertációm az innováció és a felelősség kapcsolatát vizsgálja, különös tekintettel ezek filozófiai, etikai és társadalomtörténeti vonatkozásaira. A téma komplexitása okán, elsősorban esettanulmányokon keresztül tudom bemutatni a különböző aspektusokat, ezért az elemzésem tárgyául a mesterséges intelligencia (AI) területét választottam, mert jelenleg a legdinamikusabban fejlődő innovációs irányként azonosítottam, amely kiválóan szemlélteti a technológiai fejlődés által felvetett etikai dilemmákat. A dolgozatom célja, hogy az elérhető szakmai állásfoglalások, etikai kódexek és nemzetközi ajánlások szisztematikus elemzésén keresztül feltárja az innovációs folyamatok etikai szabályozásának alapelemeit, a fő tematikus csomópontokat és a kortárs etikai elveknek való megfeleléseket, kapcsolódásokat és eltéréseket. A kutatást egy alkalmazott etikai megközelítésben végzem el, és maga a témám eredendően interdiszciplináris jellegű, mivel az innováció etikai vonatkozásainak vizsgálata a filozófia mellett más tudományágak – történettudomány, szociológia, közgazdaságtan – elméleti és módszertani bázisának integrációján alapul.

Az 1999-ben *A Creative Marginality: Innovation at the Intersections of Social Sciences* címmel megjelent könyvében a francia társadalomtudós Mattei Dogan és Robert Pahre a társadalomtudományok interdiszciplináris jellegét és a tudományterületek közötti hibridizáció szerepét vizsgálja, bemutatva hogyan járul hozzá a különböző diszciplínák közötti együttműködés a tudományos előrehaladáshoz. Kiemelik, hogy a rokon szakterületek közötti érintkezés során a szomszédos diszciplínák koncepciókat és elméleteket cserélnek egymással és a szakterületek, amelyek hosszabb ideig kölcsönhatásban állnak, gyakran innovatív hibrideket alkotnak (Dogan & Pahre, 1999). A jelenlegi tudásanyag nagysága miatt a specializáció elengedhetetlen, ez teszi lehetővé teszi a kutatók számára, hogy elődeik munkájára építsenek, mert a rekombináció értelmetlen lenne a két különálló területen korábban elért előrelépés nélkül.

„A fontos kérdések nem tartják tiszteletben a tudományágak közötti hivatalos határvonalakat” – írja Dogan (2003, p. 127) egy másik tanulmányában. A hibridizáció fogalmát, azon speciális szakterületek leírására használja, amelyek egyes diszciplínák részterületei között jönnek létre. *„A hibridizáció fogalma elsődlegesen fogalmak és elméletek kölcsönvételét és kölcsönadását jelenti”* (Dogan, 2003, p. 107) – azaz a hibridizáció a különböző tudományágak megtermékenyítő teoretikus kölcsönhatása. *„A hibridizáció az egyes tudományágak részterületeinek átfedésén alapul, a tudás új szakterületek képében jelentkező*

rekombinálódásán” (Dogan, 2003, p. 126) – a hibridizáció lényege, hogy a kutatók képesek legyenek a szakterületek közötti határok átlépésére és a releváns ismeretek integrálására, anélkül, hogy feltétlenül mélyreható jártasságot kellene szerezniük az érintett diszciplínák teljes spektrumában. A hibrid tudományok művelése rugalmas és adaptív hozzáállást igényel, amely lehetővé teszi a kutatók számára, hogy a hagyományos diszciplináris határokon átívelő módon közelítsenek kutatási problémáikhoz.

A fentiek értelmében az innovációetika eredendően egy autonóm hibrid terület és ebből következnek a dolgozatom módszertani sajátosságai is. Az alkalmazott etika a teoretikus kutatásokat alárendeli az *alkalmazhatóság* feltételeinek, ennek eredményeképpen konkrét szerepet vállal pl. szakpolitikai döntéshozatali folyamatok előkészítésében. Mindez az etika további fragmentálódását hozza magával, új határtudományok, szubdiszciplínák létrejöttét jelenti – még az innovációetika speciális területén belül is. Emiatt a diszciplináris fragmentáltság és hibriditás miatt szükséges kilépnem a filozófia és az etika szaktudományos kereteiből és az alkalmazás és az alkalmazhatóság jegyében egy tudományközi szemléletben vizsgálni a témám.

A dolgozatom II. fejezete a kulcsfogalmak – alkalmazott filozófia és etika, innováció és felelősség – konceptuális analízisét végzi el. A fogalmak pontos definiálása és kontextusfüggő használatuk tisztázása elengedhetetlen a félreértések elkerüléséhez és a világos érveléshez. A fejezet a fogalmak jelentésárnyalatait, történeti és szaktudományos kontextusát is figyelembe veszi, megalapozva a további vizsgálódást. A dolgozat III. részében áttekintem az innováció fogalmi fejlődésének filozófiatörténeti aspektusait, hogyan változott az innováció értelmezése az évszázadok során, és milyen jelentősebb gondolkodók, irányzatok járultak hozzá a koncepció fejlődéséhez. A IV. fejezet a kortárs innovációetikai diskurzus főbb témáit és irányait mutatja be, kitérve olyan aktuális kérdésekre, mint a felelős innováció vagy a felelősségetikai szemlélet jelentősége az innovációs folyamatok vizsgálatában és értékelésében. Ezt követően az V. fejezetben egy esettanulmányban vizsgálom a mesterséges intelligencia fejlődésének etikai kihívásait, a különböző etikai irányelveket és szabályozási törekvéseket, amelyek az AI fejlesztésének és alkalmazásának felelősségteljes kereteit próbálják meghatározni. A fejezet kitér a magas kockázattal bíró területek sajátosságainak az elemzésére és a felelősségetikai megközelítés relevanciájára. Végül a VI. fejezetben mutatom be az általam kidolgozott etikai döntéshozatali modellt, az felelős döntéshozatali mátrixot, amely segít felmérni az innovációk etikai dimenzióit és kiszámítani egy *reszponzibilitási indexet*, amely a felelős tervezés mértékét mutatja. Ebben a fejezetben kísérletet teszek egy átfogó, koherens innovációetikai

keretrendszer felvázolására, amely ötvözi a filozófiai belátásokat a gyakorlati alkalmazhatóság szempontjaival. Különös hangsúlyt fektetnek a felelősség és az elővigyázatosság elveire, amelyek elengedhetetlenek az innovációs folyamatok hosszú távú, társadalmilag és környezetileg fenntartható menedzseléséhez. Ez a keretrendszer ötvözi a filozófiai, etikai és kultúrtörténeti belátásokat a kortárs innovációs gyakorlatok kihívásaival, és konkrét irányelveket fogalmaz meg a felelős innováció megvalósításához.

A dolgozatomban amellet érvelek, hogy az innovációs folyamatok etikai dimenzióinak szisztematikus feltárása és az ezen alapuló normatív keretrendszer kialakítása elengedhetetlen a 21. század kihívásainak kezeléséhez. Csak egy átfogó, filozófiaiilag megalapozott és gyakorlatban alkalmazható innovációetikai megközelítés biztosíthatja, hogy a technológiai fejlődés valóban az emberi jóllét és fejlődés szolgálatába álljon, összhangban az emberi méltóság és a környezeti fenntarthatóság alapvető értékeivel.

A disszertációmban a következő kutatási kérdésekre keresem a választ:

- K₁: Milyen definíciós problémák merülnek fel az innováció és felelősség kontextusában, és hol van ezeknek a fogalmaknak a helye a kortárs filozófiai diskurzusban?
- K₂: Miként járul hozzá az innováció fogalom történeti fejlődésének vizsgálata az innováció által generált kortárs etikai problémák tárgyalásához?
- K₃: Milyen helyet foglal el és milyen súllyal jelenik meg a felelősségetika és az innovációetika a kortárs tudományos diskurzusban?
- K₄: Milyen filozófiai és etikai elméletek játszanak kulcsszerepet a felelős innováció, mint praxis kereteinek kialakításában?
- K₅: Melyek azok az etikai princípiumok és követelmények, amelyek elengedhetetlenek egy konzisztens felelősségetikai modell kialakításához az innovációs folyamatok etikai szempontú megfeleléségének értékelésére? Milyen módszertani lehetőségek és korlátok merülnek fel egy ilyen modell létrehozása során?
- K₆: Milyen etikai dilemmákat és társadalmi kihívásokat generál a mesterséges intelligencia fejlődése és széleskörű alkalmazása? Hogyan képes a felelős innováció hozzájárulni a mesterséges intelligencia fejlesztésének és alkalmazásának etikai szempontú szabályozásához?

Az előzőekben megfogalmazott kutatási kérdések alapján két fő hipotézist állítottam fel:

H₁: Az erkölcsi elvek és normák következetes alkalmazását az innovációs folyamatok tervezése, megvalósítása és értékelése során akadályozza egy egységes, koherens és széles körben elfogadott etikai elméleti keret hiánya, mivel a különböző filozófiai és etikai megközelítések eltérő gyakorlati iránymutatásokat kínálnak

H₂: A felelősségetika alapelvei integrálhatók egy értékelési modellbe, amely képes mérni az innovációs folyamatok etikai megfelelőségét, és ezen keresztül elősegíteni a felelős innováció gyakorlatának elterjedését a technológiai és más tudományos fejlesztések terén.

Disszertációmmal szeretnék hozzájárulni az innovációetika elméleti alapjainak megszilárdításához, ugyanakkor reményeim szerint a munkám gyakorlati relevanciával is bír, hiszen az innovációs folyamatok etikai szabályozása kulcsfontosságú a technológiai fejlődés társadalmilag felelős és fenntartható alakításához. A kutatás eredményei hasznosak lehetnek a döntéshozók, a kutatók és a fejlesztők számára, segítve őket abban, hogy az innováció valóban az emberiség javát szolgálja, miközben minimalizálja a potenciális kockázatokat és negatív következményeket.

A felelős innováció koncepciójának előtérbe helyezésével a disszertációban hangsúlyozni szeretném, hogy a technológiai fejlődésnek mindig az emberiség javát kell szolgálnia, miközben minimalizálja a potenciális kockázatokat és negatív következményeket. Az innováció csak akkor lehet valóban sikeres és értékteremtő, ha etikai és társadalmi szempontokat is figyelembe vesz, és harmonikusan illeszkedik az emberi kultúra és értékek rendszerébe.

1.1 Módszertani megjegyzések

Dolgozatomban a Pécsi Tudományegyetem Filozófia Doktori Iskolájának szabályzatával összhangban lévő irányelveket tartalmazó APA 7 (American Psychological Association) formátum és stílusútmutatóját követem, amely a társadalom és bölcsészettudományok területén széles körben alkalmazott hivatkozási elveket ad meg. Szó szerinti idézet esetében szerepeltetem a pontos oldalszámot – kivéve, ha olyan pl. online hozzáférésű forrásból dolgoztam, ahol ez nem elérhető. Ebben az esetben a szerző neve és a megjelenés évszáma után az s.p. (sine pagina – oldalszám nélkül) jelölés szerepel. Amennyiben nincs az adott szövegnek korábbi magyar fordítása, a dolgozat törzsszövegében az idézeteket eredeti nyelven vagy az elérhető tudományosan széles körben elfogadott angol nyelvű fordításban adom meg, és a

lábjegyzetben közlöm az értelmezést segítő saját fordításom. A lábjegyzetben – az APA ajánlásait követve – csak a legszükségesebb információkat jelenítem meg, a szövegkohézió és az olvashatóság miatt a dolgozat lényegi információit a törzsszöveg tartalmazza.

A PTE Filozófia Doktori iskola hatályos szabályzatának megfelelően, a doktori dolgozatomba beépülnek a Kooperatív Doktori ösztöndíjas időszakom (2020-24) alatt megjelent publikációim egyes részletei is: ezek több ponton átalakításokon, revíziókon estek át, sok esetben újrastrukturálódtak és kiegészítést kaptak. A négyéves kutatásom egy-egy szakaszát lezáró tanulmányok szerves részét képezik és fontos állomásait jelentik a kutatási tervemben meghatározott munkafolyamatnak. A megállapításaim a disszertáció elkészítésének folyamatában egy kritikai szűrőn mentek át, több esetben korrigáltam azokat. A felsorolt művek a szövegben megfelelően hivatkoztak és az irodalomjegyzékben is tételesen felsorolásra kerültek, pontos bibliográfiai forrásjelöléssel:

- Szilágyi, T. (2021). A felelősség társadalma. *Kultúratudományi Szemle*, 3(4), 38-45.
- Szilágyi, T. (2022). A felelősség kérdése a kortárs etikai diskurzusban. In G. Szécsi, J. Boros & Zs. Nemeskéri (Eds.), *XXI. századi narratívák: Tanulmányok a filozófia, művészet és gazdaság metszéspontján* (pp. 111-122). Belvedere Meridionale.
- Szilágyi, T. (2023). A felelősség fogalma Richard Mckeon filozófiai gondolkodásában. In E. Daróczy & S. Laczkó (Eds.), *A felelősség* (pp. 385-389). Státus Könyvkiadó, Pro Philosophia Szegediensi Alapítvány, Magyar Filozófiai Társaság.
- Szilágyi, T. (2024a). Innováció vs. fenntarthatóság: A kortárs tudományos gyakorlat etikai dilemmái. In F. Dudok (Ed.), *Környezeti nevelés a 21. században: Tapasztalatok és jó gyakorlatok* (pp. 142-150). Pécsi Tudományegyetem.
- Szilágyi, T. (2024b). A mesterséges intelligencia alkalmazásának etikai vonatkozásai az oktatási innovációban. In T. K. Fodoroné (Ed.), *Felsőoktatási Lifelong Learning a képesség-fejlesztés, a kultúraközvetítés és fenntarthatóság tükrében. A hatékony és innovatív tudástranszferek lokális/regionális és globális trendjei* (pp. 209-214). MELearn Egyesület.

Befogadott, megjelenés alatt:

- Szilágyi, T. (2024c). *A felelős innováció szerepe a nem szándékolt következmények mérséklésében*. In Innováció és társadalmi felelősség – tanulmánykötet, Pécsi Tudományegyetem (várható megjelenés: 2024 ősz)
- Szilágyi, T. (2024d). *Az innováció etikai kihívásai az emberi erőforrás fejlesztésben*. In Informális tanulás és új tudásformák a 21. században – tanulmánykötet, Széchenyi István Egyetem (várható megjelenés: 2024 ősz)
- Szilágyi, T. (2024e). *Etikai megfontolások az innováció biztonságpolitikára gyakorolt hatásaival kapcsolatban*. In Scientia et Securitas (várható megjelenés: 2024 ősz)
- Szilágyi, T. (2024f). *A közösségi médiafelületek mint pseudo-nyilvános terek: a felügyeleti kapitalizmus etikai dilemmái*. In Média, edukáció, felelősség – tanulmánykötet, SZTE-JGYPK (várható megjelenés: 2024 ősz)

II. Fogalmi konceptualizáció

A dolgozatom első fejezetében a későbbiekben vizsgálandó fogalmak – az innováció és a felelősség –, illetve a módszertani háttér – az alkalmazott filozófia és etika –, egyfajta konceptuális analízisét végzem el, azzal a céllal, hogy pontos kereteket szabjak a fogalmak jelentésének tisztázásával, definíciójuk megadásával és a kontextusfüggő használatukkal kapcsolatosan. Ezen elemzés célja, hogy egyértelműen kijelölje a metodológiai határokat, meghatározza és elmagyarázza a kulcsfogalmakat mielőtt részletesen tárgyalná őket, ezáltal biztosítva a pontos és világos érvelést, valamint elkerülve a félreértéseket a tudományos diskurzusban. Igyekszem tekintetbe venni az értelmezések legszélesebb spektrumát, az egyes fogalmak jelentésárnyalatait, történeti és szaktudományos kontextusát, kialakítva azokat a kritériumokat, amely alapján alkalmazni fogom őket. Mindezek mellett nem végeztem el a fogalmak komplex nyelvi vagy történeti analízisét (a dolgozat további fejezeteiben részletesen tárgyalom ezeket), ez a fejezet a konceptualizációt célozza és a tudományos kontextust vázolja fel.

2.1 Az alkalmazott filozófia, mint teoretikus keret

Jelen dolgozat – tematikus sajátosságai révén –, elsődlegesen az úgynevezett alkalmazott filozófia, közelebbről az alkalmazott etika szaktudományos apparátusával és elméleti bázisán kívánja tárgyalni fő kérdéseit. Mivel napjainkban az alkalmazott filozófiai/alkalmazott etikai kutatások már a filozófiai gondolkodás fősodrában vannak és habár nem tűntek el, de egyre kevésbé a korábbi definíciós viták és diszciplináris lehatárolások körül zajlik a szakmai diskurzus (Callicott, 1990; LaFollette, 2003; Warren, 1992), nem látom szükségesnek részletes érvelésbe bocsátkozni a választott megközelítés tudományos relevanciájának bizonyítására. Azonban szükségesnek érzem azon keretek pontos meghatározását, amelyekben a következőkben a megállapításaim teszem, így esetünkben nem tudok eltekinteni az alkalmazott filozófia és az alkalmazott etika általam használt értelmezésének megadásától, elkerülve a hibás definíciókból és interpretációkból fakadó félreértéseket.

A komplex gyakorlati kérdések, a valós problémák értelmezésében-kezelésében a filozófiai módszerek és teóriák fontos szerepet töltenek be napjainkban, nem csupán az elméletalkotás, de a cselekvés tényleges modelljeinek meghatározásában is. Maga az alkalmazott filozófia elnevezés a hatvanas évektől kezdve kap nagyobb figyelmet, a korszak társadalmi, politikai és tudományos folyamataira reflektálva. Leslie Stevenson az 1970-ben írt *Applied philosophy* című tanulmányában veti fel egy olyan filozófia szükségességét, amely

releváns válaszokat képes kínálni „*a mindennapi élet fontos kérdéseire*” (Stevenson, 1970, p. 258). A hagyományos – általa *tisztának* (pure) nevezett – filozófia mellett, az *alkalmazás* kereteinek tisztázását is fontosnak látja, meghatározva, hogy mit ért ez alatt:

„Under pure philosophy, I intend to include most of what is now studied as philosophy in the universities, viz. mathematical and philosophical logic, metaphysics, epistemology, philosophy of mind, philosophy of science, and most of the questions now discussed by professional philosophers about ethics, politics, and aesthetics (e.g. the validity of the fact-value distinction). As examples of applied philosophy I suggest the rational discussion of particular controversial moral questions, such as sexual morality, the Catholic ban on contraception, the use of hallucinogenic drugs, abortion, euthanasia, eugenics, the definition of death, and many other medico-ethical-legal problems raised or soon to be raised by the coming “biological revolution”; also certain aspects of various difficult social and political problems, such as educational policy (comprehensive schools? religious education?), the need for public participation in planning (Do people know what they want twenty years from now, and is it identical with what they need? How can many different pressures result in a sensible and just decision?), world economic development (Do the richer countries have a duty to help the poorer?)”¹ (Stevenson, 1970, pp. 258-259).

A terület első jelentős tudományos szervezete, a Society for Applied Philosophy (SAP) 1982-ben jött létre azzal a céllal, hogy az alkalmazott filozófia területén folytatott gyakorlatorientált tanulmányokat és kutatásokat támogassa. Ez a törekvés azon a felismerésen alapult, hogy a filozófia sajátos kritikai és elemző módszerei értékes eszközök lehetnek számos közéleti kérdés mélyebb megértéséhez és megvitatásához. Például az alkalmazott filozófia relevanciája széles körben megmutatkozik, többek között olyan területeken, mint a jogtudomány, politika, közgazdaságtan, természettudományok, technológia, egészségügy és oktatás (SAP, 2024). A társaság adja ki az alkalmazott filozófia heterogén témáinak szentelt nemzetközi tudományos szaklapot *Journal of Applied Philosophy*-t, amely sikeresen tematizálja az elmúlt évtizedekben a területen zajló szakmai párbeszédet és vitákat.

¹ „*A tiszta filozófia alá kívánom sorolni mindazt, amit ma filozófiaként tanulnak az egyetemeken, azaz a matematikai és filozófiai logikát, a metafizikát, az ismeretelméletet, az elmefilozófiát, a tudományfilozófiát, valamint a legtöbb olyan kérdést, amelyet ma a hivatásos filozófusok az etikával, a politikával és az esztétikával kapcsolatban tárgyalnak (pl. a tény-érték megkülönböztetés érvényességét). Az alkalmazott filozófia példaként javaslom egyes vitatott erkölcsi kérdések racionális megvitatását, mint például a szexuális erkölcs, a fogamzásgátlás katolikus tilalma, a hallucinogén drogok használata, az abortusz, az eutanázia, az eugenika, a halál definíciója és számos más orvos-etikai- etikai-jogi problémák, amelyeket a közelgő 'biológiai forradalom' vet fel vagy fog hamarosan felvetni; továbbá különböző nehéz társadalmi és politikai problémák egyes aspektusait, mint például az oktatáspolitikát (általános iskoláztatás? vallásos oktatás?), a lakosság részvételének szükségessége a tervezésben (Tudják-e az emberek, hogy mit akarnak húsz év múlva, és ez azonos-e azzal, amire szükségük van? Hogyan vezethet sok különböző nyomás ésszerű és igazságos döntéshez?), világméretű gazdasági fejlődés (Kötelesek-e a gazdagabb országok segíteni a szegényebbeket?)” (saját fordítás)*

Gimler és munkatársai (2023, p. 108) *The Varieties of Applied Philosophy: Introduction* címen megjelent tanulmánya, az alkalmazott filozófia három jellegzetességét emeli ki: [1] sokféle formában létezik, ugyanolyan széles spektrumú, mint a tiszta filozófia és a módszerei között megtalálható az analitikus filozófia, a kritikai elmélet, a hermeneutika éppúgy, mint a fenomenológia; [2] még fejlődésben van, új formái alakulhatnak ki a jövőben; [3] bár különböző filozófiai hagyományokból erednek, az alkalmazott filozófiák közös vonása, hogy szkeptikusak a filozófiai rendszerek magyarázó erejével szemben. Az általuk megadott definíció az alkalmazott filozófia szerepét a SAP küldetésnyilatkozatában leírtakhoz hasonlóan, a filozófián (mint tudományon) kívüli területek vizsgálatában látja, határozott gyakorlati fókusszal: „[...] *applied philosophy is a branch of philosophy that applies traditional philosophical concepts, theories, and methods to problems originating from situations that arise in practices outside academia itself*”² (Gimler et al., 2023, p. 108).

Karikó Sándor 2002-es *Mi az alkalmazott filozófia?* címet viselő tanulmányában felhívja a figyelmet, hogy „[...] *az alkalmazott filozófia egyáltalán nem a filozófia alkalmazását, mintegy az elméletnek a gyakorlatra való lefordítását jelenti*” (Karikó, 2002, p. 42), hanem szerinte az alkalmazott filozófia egy összetett megközelítés, amely túlmutat az elvont filozófiai koncepciók egyszerű gyakorlati alkalmazásán, nem tekinthető pusztán az absztrakt filozófia leegyszerűsített vagy gyakorlatiasabb verziójának, sem pedig annak igazolására vagy hasznosságának bizonyítására szolgáló eszköznek. Az alkalmazott filozófia nem redukálható pusztán technikai megoldások vagy módszertani útmutatók gyűjteményére. Ehelyett egy önálló, komplex diszciplínaként kell látnunk, amely sajátos megközelítéssel és eszköztárral rendelkezik a gyakorlati problémák filozófiai vizsgálatára és megoldására.

Király V. István az alkalmazott filozófia lényegét abban látja, hogy képes kilépni a hagyományos filozófiai keretből, és aktuális kérdéseket vizsgál úgy, hogy nem csak reflektál rájuk, hanem aktívan részt vesz azok alakításában és felépítésében is. Az alkalmazott filozófia tehát nem csupán meglévő filozófiai kereteket alkalmaz, hanem maga is hozzájárul a vizsgált kérdések formálásához és strukturálásához (Király, 2002, p. 23). Király egy másik írásában felhívja a figyelmet arra, hogy az alkalmazott filozófia megalapítói maguk is kétségesnek látták valamiféle egységes elméletét kidolgozni az alkalmazott filozófiának, mivel például az esztétikával szemben nem a filozófia valamely ágazatként kell tekintenünk rá „[...] *hanem*

² „*az alkalmazott filozófia a filozófia egyik ága, amely a hagyományos filozófiai fogalmakat, elméleteket és módszereket alkalmazza olyan problémákra, amelyek a tudományos életen kívüli gyakorlatban felmerülő helyzetekből erednek.*” (saját fordítás)

inkább a filozófusoknak azon erőfeszítése, hogy a napi társadalmi, illetve gyakorlati problémákra ügyeljenek” (Király, 2002, p. 23).

Tóth I. János rámutat az alkalmazott filozófia által vizsgált területek heterogén voltára, amely *„nagyban megnehezíti az alkalmazott filozófia és etika egységes tárgyalását”* (Tóth, 2002, p. 60). Egyetért abban, hogy az alkalmazott filozófia a mindennapi élet szempontjából fontos, filozófiai relevanciával is bíró kérdésekre igyekszik választ adni, konkrét problémákból kiindulva.

Szécsi Gábor megállapítja, hogy a hagyományos, egységes filozófiai nyelv felbomlott, ahogy különböző filozófiai iskolák elutasították vagy módosították a hagyományos terminológiát és grammatikát. Ez a filozófia más tudományágakkal való szorosabb integrációjához vezetett és az alkalmazott filozófia ennek a változásnak az eredménye. Az alkalmazott kutatás és a filozófiai nyelv konvergenciája lehetővé tette a filozófiai fogalmak használatát olyan területeken, mint a neurobiológia és a pszichológia, miközben a filozófusok is mindinkább aktuális, gyakorlati kérdések megvitatása felé fordultak. A filozófiai nyelv alkalmazhatósága című tanulmányban az alkalmazott filozófia legkomplexebb magyar nyelvű definícióját is megtaláljuk:

„[...] alkalmazott filozófia alatt olyan filozófiai vizsgálódásokat értünk, amelyek kritikai és analitikus módszer alkalmazásával közvetlenül az emberi gyakorlat valamely elemére irányulnak. Így terjedhetnek ki az alkalmazott filozófiai kutatások a környezeti és orvosi etika kérdéseire, a tudományos és technológiai változások társadalmi következményeire, az oktatás, jog és gazdaság filozófiai és etikai vonatkozásaira” (Szécsi, 2002, pp. 33-34).

A fentiek alapján látható, hogy amikor alkalmazott filozófiáról beszélünk, akkor nem egy szűk definitív keretek közé szorított megközelítést tárgyalunk. Kasper Lippert-Rasmussen az alkalmazott filozófia természetéről szóló tanulmányában annak hét eltérő megközelítését határozza meg: [1] a relevancia központú megközelítés (the relevance conception) szerint, még ha nem is képes feltárni és megválaszolni a filozófia a mindennapi élet fontos kérdéseit, azonban ezen kérdések szempontjából érvényesnek kell lennie. [2] Az egyedi sajátosságokat középpontba helyező (the specificity conception) megközelítés alapján a filozófia csak akkor használható hatékonyan, ha egy adott filozófiai ágon belül (pl. metafizika, ismeretelmélet, etika) konkrét kérdést vizsgál. Alapelveket alkot, majd ezeket alkalmazza és elemzi hatásukat nem filozófiai területeken. [3] A gyakorlati koncepció (the practical conception) – amely konkrét kérdésekre ad válaszokat például, hogy *mit kellene tennünk*. [4] Az aktivista felfogás (the activist conception) olyan megközelítés, amelyet az a törekvés motivál, hogy bizonyos

hatást gyakoroljon a világra. Az alkalmazott filozófia célja ebben a megközelítésben a világra gyakorolt konkrét hatás, legyen az nevelés, felvilágosítás vagy építés. Bár a *hagyományos* (Stevenson-i értelemben: tiszta) filozófia is befolyásolhatja a világot, az alkalmazott filozófia kifejezetten erre törekszik. Ez a megközelítés a tudás és igazság keresése helyett az oksági hatásra összpontosít, ami módszertani változásokat is eredményezhet. [5] A módszertani megközelítés (the methodology conception) középpontjában az a törekvés áll, hogy filozófiai módszereket alkalmazzon a szűk filozófiai problémákon kívüli kérdések feltárására és elemzésére. [6] Az empirikus tények koncepciója (the empirical facts conception) egy jelentős mértékben empirikus bizonyítékok által támogatott filozófiai megközelítés, amelyben fontos szerepet kap az empirikus tudományok által szolgáltatott tények és bizonyítékok használata, egy felvállaltan interdiszciplináris megközelítésben. És végül a [7] közönség központú koncepció (the audience conception) amely egy nem a szakmának, nem a filozófiai szempontból képzett közönségnek szánt filozófia (Lippert-Rasmussen 2016, pp. 7-15).

A szerző rámutat, hogy az alkalmazott filozófia és a hagyományosnak tekintett filozófia közötti különbségtétel nem éles határvonalak mentén történik meg, sokkal inkább fokozati kérdés. Alapvetően módszertani megközelítések választásának a kérdése, hogy melyik felfogás mentén dolgozunk. A fenti koncepciók együttes megléte nem feltétele annak, hogy egy filozófiát alkalmazottnak minősítsünk, ahogy nem is zárja ki, hogy a különböző koncepciók együttes jelenléte kirekessze a megközelítésünk az alkalmazott filozófia köréből. Lippert-Rasmussen az alkalmazott filozófia tágabb értelmezése mellett érvel, ami túlmutat az alkalmazott etikán, amelyet megalapozott aldiszciplínaként határoz meg, utalva növekvő jelentőségére az alkalmazott filozófián belül (szembeállítva más szubdiszciplínákkal, amelyek nem szükségszerűen jelennek meg alkalmazott formában, mint például az esztétika). Hasonlóképp utal Király (2004) is az etikai megközelítés növekvő súlyára az alkalmazott filozófiában:

„a magukat „alkalmazottnak” nevező filozófiai kutatásokban ténylegesen és talán túlságosan is nagy szerepet kap az „etika”...Azért mondom, hogy talán „túlságosan is”, mert így szerintem fennáll a veszély, hogy - ha nem más miatt, hát a diszciplináris beidegződések folytán - az alkalmazott filozófiai kutatásokat - és főként azok teoretikus-filozófiai horderejét - eleve és filozófiai értelemben is pusztán a gyakorlati, azaz a praktikus filozófiára szűkítsük” (Király, 2004, p. 19).

Az alkalmazott filozófiát jelen dolgozat keretében én az alábbi módon értelmezem: Az alkalmazott filozófia egy *komplex, rugalmas és változásorientált* tudományos megközelítésmód, amely a hagyományos filozófiai fogalmakat, elméleteket és módszereket

alkalmazza *gyakorlati problémák és helyzetek kutatására, elemzésére és megoldására*. Aktívan részt vesz az általa vizsgált kérdések alakításában és strukturálásában, interdiszciplináris jellegéből adódóan *más tudományágak módszereit és eredményeit integrálva* igyekszik választ adni a társadalmi, a tudományos vagy a mindennapi élet területét érintő kérdésekre. A tematikus heterogenitásából adódóan egyaránt vizsgál olyan területeket, mint a tudományos és technológiai változások társadalmi következményei, az oktatás, a jog és a gazdaság filozófiai bölcséleti aspektusai – diszciplináris nyitottságából adódóan *folyamatosan bővítve területeit*.

2.2 Az alkalmazott etika sajátosságai

Mivel jelen dolgozatban elsődlegesen az alkalmazott etika területén végzem a vizsgálataim, olyan fogalmakat – például a felelősség – tárgyalva, amelyek etikai relációi az elsődlegesek, fontosnak tartom az alkalmazott etika lehatárolását és azon definíciós határvonalak kijelölését, amelyek a kutatásom kereteit egyértelműen megadják. Az alkalmazott filozófia tágabb jelentéstartamához viszonyítva kell meghatároznom a választott módszertani és teoretikus kereteket, amelyekben a következőkben dolgozni fogok.

Amennyiben a fenti elméleteket és definíciós kísérleteket elfogadjuk és az alkalmazott filozófiára nem szubdiszciplínaként, hanem egy dinamikusan változó, a maga határait kereső tudományos megközelítésként tekintünk, akkor ennek tükrében kell az alkalmazott etikát is meghatároznunk.

Ahogy láttuk, az alkalmazott filozófia számos alterületet – ismeretelmélet, logika, metafizika – magába foglal: az alkalmazott etika ezek közül az egyik, amely célzottan erkölcsi kérdésekkel, dilemmákkal, az emberi cselekvés erkölcsi mozgatórugóival és határaival foglalkozik elemző módon, gyakorlati kontextusban. Az alkalmazott etika fókuszában álló területek sok esetben szaktikák: bioetika, környezetetika, üzleti etika, orvosi etika – fókusza ezért az alkalmazott filozófiánál szükségszerűen jóval szűkebb, specifikusabb. Módszertanilag az alkalmazott filozófiához hasonlóan interdiszciplináris jellegű, számos szakmai területtel kell együttműködnie azok sajátosságait ismerve és alkalmazva.

Brenda Almond a Society for Applied Philosophy társalapítója szerint az alkalmazott etika lényege, az etikai szempontok és elvek *felhasználása* különböző helyzetek *kezelésére és megoldására*, amely magában foglalja erkölcsi értékek, alapelvek és eszmék *alkalmazását* mind egyéni, mind társadalmi szintű szabályozások és gyakorlatok vizsgálatában és az ezzel kapcsolatos döntéshozatalban (Almond, 1993, p. 33). A cél az, hogy etikai alapon értékeljük ezeket, és ennek eredményeként támogassuk vagy vessük el őket. Ez a megközelítés egy olyan

döntéshozatali modellt képvisel, amely kifejezetten az etikai megfontolásokra összpontosít, ezáltal különbözik más döntéshozatali módszerektől, amelyek például a legelőnyösebb opciót keresik, vagy kizárólag az egyéni érdekekre koncentrálnak. Almond szerint az alkalmazott etika elfogadásával, elutasítjuk azt a nézetet, hogy a filozófia csak elemezni és tisztázni tudja az erkölcsi problémákat, de nem képes felvállalni a rájuk adandó válaszok keresését. Vagyis ebben a megközelítésben benne van, hogy az alkalmazott etika az emberi élet különböző területein az erkölcsileg helyes cselekvési irányok meghatározására és befolyásolására irányuló filozófiai módszerekkel tett kísérleteket is magában foglalja. Az alkalmazott etika tehát egy olyan eszköz, amely segít morális alapon mérlegelni és értékelni különböző cselekvési lehetőségeket és gyakorlatokat, iránymutatást adva azok alkalmazására (Almond, 1995).

Beauchamp rámutat arra, hogy az alkalmazott etika meghatározása, mint az általános etikai elméletek konkrét erkölcsi problémákra történő alkalmazása, túl szűknek bizonyulhat. Problematikusnak látja, hogy egyes filozófusok gyakran úgy tekintenek az alkalmazott etikára, mint egy általános erkölcsi normák vagy elméletek gyakorlati problémákra való alkalmazására, mivel a gyakorlati szabályozásokhoz nem vezet egyenes út a teória vagy az általános elvek felől közelítve. Szakadék húzódik a teória és a praxis között, amelyet adatok, szervezeti tapasztalatok és egyedi helyzetek elemzésének felhasználásával lehet áthidalni. Beauchamp kritizálja azt a felfogást, amely szerint az alkalmazott etika szinonimája a szakmai etikának. Az alkalmazott etika által vizsgált területek sokkal több kérdést ölelnek fel annál, hogy specifikus szakterületi etikaként tekintsünk rá. Az olyan kérdések, mint a háborúk, érdekkonfliktusok, állatkísérletek, adatvédelem - messze túlmutatnak a szakmai viselkedés szabályozásain, és az alkalmazott etika tartalmi és módszertani megközelítését is érintik, így szélesebb kontextusban kell azokat vizsgálni (Beauchamp, 2003, pp. 2-3).

Bedeau az alkalmazott etikai módszerei kapcsán kiemeli a gyakorlati mérlegelés (practical reasoning) jelentőségét, amely során etikai elveket és elméleteket használunk valós helyzetek erkölcsi értékelésére és döntéseink igazolására. Ez magában foglalja különböző etikai szempontok és érvek gondos számbavételét, érvek felsorakoztatását, hogy észszerű etikai következtetésekre jussunk:

„applied ethics is the application of ethical considerations—reasons, principles, values, ideals—to any policy or practice—personal or social—for the purpose of evaluating (and thus endorsing or rejecting) that policy or practice on ethical grounds. Thus, applied ethics is the branch of practical reasoning in which ethical

*(as opposed to prudential or selfish) considerations are employed*³ (Bedeau, 2011, pp. 80-81).

Ez azért különösen fontos az alkalmazott etika területén, mivel ahogy az előzőekben láttuk, az általános normák és elméletek gyakran csak kiindulópontként szolgálnak, és szükséges a konkrét helyzetekhez illő cselekvési modellek és normák kidolgozása.

Az alkalmazott filozófiához hasonlóan az alkalmazott etika sem az elméleti etika gyakorlati alkalmazása. Sokkal inkább egy összetett, kétirányú folyamatról van szó, ahol az elmélet és a gyakorlat kölcsönösen hatnak egymásra, akként, hogy az etikai elméletek inkább iránymutatást adnak és segítenek a problémák etikai aspektusainak feltárásában és azonosításában, mintsem, hogy kész válaszokat kínálnának. A különböző etikai megközelítések eltérő szempontokat kínálnak, ezzel gazdagítva az elemzés spektrumát és a megoldások kidolgozásának lehetőségeit.

Az alkalmazott etikában kulcsfontosságú a megfelelő normatív érvelés és a racionális vita. Végso soron az alkalmazott etika nem csupán elméletek alkalmazása, hanem egy dinamikus folyamat, ahol az elmélet és a gyakorlat kölcsönösen formálják egymást úgy, hogy az etikai elméletek gyakran konkrét esetek tanulmányozásából fejlődnek ki (Van de Poel & Royakkers, 2011, p. 106).

Beauchamp és Childress leszögezik a Principles of biomedical ethics című orvosetikai alapl műük bevezetésében, hogy az alkalmazott etika és az etikai elméletek kapcsolata összetett, az elméletek nem szolgálnak kész válaszokkal, hanem leginkább csak segítenek feltárni egy helyzet etikai aspektusait. A különböző megközelítések eltérő szempontokat emelnek ki, és a normatív érvelés (amely során erkölcsi ítéleteket és etikai döntéseket indokolunk meg) helyessége kulcsfontosságú a racionális etikai vitákhoz, amelyek nélkülözhetetlenek az alkalmazott etika területén:

„Practical ethics-used here as synonymous with applied ethics, and by contrast to theoretical ethics-employs general concepts and norms to address particular problems. The term practical refers to the use of norms and other moral resources in deliberating about problems, practices, and policies in professions, institutions, and public policy. Often no straightforward movement from general norms, principles, precedents, or theories to particular judgments is possible. General

³ „[...] az alkalmazott etika etikai megfontolások - okok, elvek, értékek, eszmények - alkalmazása bármely eljárásra, szabályozásra vagy gyakorlatra – legyen az személyes vagy társadalmi - azzal a céllal, hogy etikai alapon értékeljük (és így támogassuk vagy elutasítsuk) azt. Az alkalmazott etika tehát a gyakorlati megfontolás azon modellje, amelyben etikai mérlegelést végzünk (szemben a legelőnyösebbet választó vagy önös döntésekkel)” (saját fordítás)

*norms are usually only starting points for the development of norms of conduct suitable for specific contexts*⁴ (Beauchamp & Childress, 1979, p. 2).

Beauchamp és Childress nevéhez kötődik a (szerzők akkori akadémiai affiliációja miatt, a később Georgetown-i alapelveknek vagy Georgetown mantrának is nevezett) négy alapelv, amely a kortárs orvosi és bioetikát jelentős mértékben befolyásolta és napjainkban is az egyik legdominánsabb megközelítésként ismert. Az általuk megfogalmazott négy alapelv a korábban *Belmont Report* (1979) címen megjelent dokumentumban közölt három alapelv [1] az autonómia tiszteletének; [2] a jótékonyosság és az [3] az igazságosság elvének kiegészítése [4] a nem-ártás elvével. Mindezen alapelvek megfogalmazásakor a fő céljuk nem egy átfogó etikai elmélet megalkotása volt, hanem sokkal inkább egy olyan konszenzuális, számos etikai-vallási rendszerből, társadalmi gyakorlatból levezethető, azokkal összhangban álló gyakorlati megközelítés megalkotása, ami biztosítja azokat az átfogó normákat, amelyek az orvostudomány területén meghatározhatják az etikus cselekvés kereteit.

Ezek az elvek jól mutatják az alkalmazott etika korlátait az emberi viselkedés és cselekvés befolyásolásában, a szabályozó / szabályalkotó lehetőségeinek korlátait felmutatva. Az elvek általános és rugalmasabb erkölcsi iránymutatások, míg a szabályok konkrétabb és merevebb előírások, amelyek az elvekre épülnek és azokat gyakorlati útmutatásokká alakítják specifikus helyzetekben (ezáltal jóval részletesebbek). Ezen a területen az alkalmazott etika nem tud előíró-szabályalkotó szerepkört felvállalni, mivel egy alapvetően pluralizmusra épülő, heterodox és interdiszciplináris szakterületen nem bír autoritással, csak interszubjektív elvek alapján képes kereteket adni.

Mureşan és Majima is az alkalmazott etika, mint *módszer* felől közelíti meg a tárgyat:

*„Today applied ethics refers to a large variety of methods capable of critically examining and assessing the moral problems of various professions, technologies, public policies etc., including (amongst others) conceptual analysis, phenomenology, casuistry, narrativism, reflective equilibrium, specification, weighing of reasons, ethical audit, ethics management, ethical matrix etc.”*⁵ (Muresan & Majima, 2013, p. 11).

⁴ „A gyakorlati etika - amelyet itt az alkalmazott etika szinonimájaként használunk, és az elméleti etikával szembeállítva - általános fogalmakat és normákat alkalmaz konkrét problémák megoldására. A gyakorlati kifejezés a normák és más erkölcsi erőforrások használatára utal az egyes szakmák, intézmények és a közpolitika problémáinak, gyakorlatainak és politikáinak megvitatása során. Gyakran nem lehetséges az általános normákból, elvekből, precedensekből vagy elméletekből egyértelmű ítélet levonása. Az általános normák általában csak kiindulópontot jelentenek a konkrét kontextusban megfelelő magatartási normák kialakításához.” (saját fordítás)

⁵ „Ma az alkalmazott etika számos olyan módszerre utal, amelyek alkalmasak a különböző szakmai területek, technológiák, közpolitikák erkölcsi problémáinak kritikus vizsgálatára és értékelésére, beleértve (többek között) a

Kiemelik, hogy az alkalmazott etika területe az évek során egyre specializáltabbá vált, a kutatások elkülönült területeken, főként kutatóközpontokban és magáncégekben folytak, igazodva az ügyfelek igényeihez. Az alkalmazott jelző nem elsősorban etikai elméletek alkalmazására utal, hanem inkább egyfajta *szabad piacra* az etikai eszközök terén, amelyeket az etikai kérdésekkel szakmai szinten foglalkozók és az egyes területen érintett alkalmazottak egyaránt használnak a cégekben, az egyetemeken vagy éppen az egészségügyi intézményekben, kórházakban. Ennek eredményeként jelentek meg olyan új szakmák, mint az etikai tisztviselő, etikai tanácsadó és etikai auditor, valamint kialakult egyfajta *etikai infrastruktúra* is, amely az alkalmazott etikát a modern menedzsment szerves részévé tette számos szervezetben.

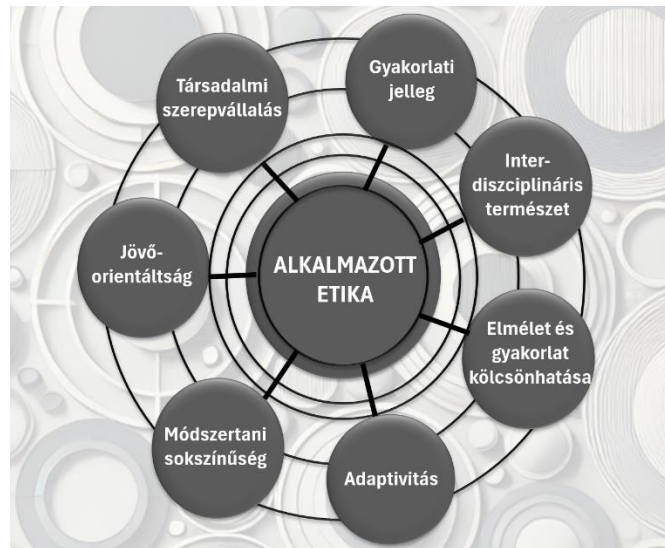
Ahogy az előzőekben már Beauchamp kapcsán utaltam rá, az alkalmazott etika több mint a szaktetikák összessége. Az alkalmazott etika az alkalmazott filozófia tematikus heterogenitását örökölve vizsgálja azokat az erkölcsi kérdéseket, amelyeket a dinamikusan átalakuló társadalmi-politikai és gazdasági környezet hívott létre. Alkalmazott filozófiaként reflexív, de egyúttal törekszik a gyakorlati dimenzióban normaadás igényével is fellépni.

A módszerek kapcsán irányadónak tekintem Richard Rorty 2005 augusztusában a Society of Business Ethics éves találkozáján elmondott beszédében megfogalmazott gondolatait, amelyben az akadémikus filozófiai megközelítés korlátozott jelentőségéről szól az alkalmazott etikában. Szerinte az erkölcsi meggyőződések igazolása mindig egy adott közönséghez és kulturális kontextushoz kötött, nem pedig abszolút érvényű. Azt javasolja, hogy az üzleti etika (mint *alkalmazott ágazati etika*) inkább a narratívákra, esettanulmányokra és kontextus-specifikus irányelvekre összpontosítson az elvont erkölcsi elméletek helyett. Véleménye szerint a filozófiai megalapozottság nem feltétlenül relevánsabb az alkalmazott etika számára, mint más tudományterületek, például a történelem, antropológia vagy az irodalom. Az üzleti etikai döntéshozatalban, is egy pragmatista megközelítést javasol, amely a jobb társadalmi lehetőségek elképzelésére összpontosít az egyetemes erkölcsi elvek keresése helyett. Végül Rorty amellett érvel, hogy az üzleti etikai közösségnek nyitottnak kellene lennie a különböző tudományterületek képviselőire, nem csak a filozófiai szakértelemre (Rorty, 2006).

Richard Rorty gondolatait követve fontos hangsúlyozni, hogy az alkalmazott etika általam használt értelmében sem kizárólag a filozófia területéről merít, hanem nyitott más tudományágak (pl. a szociológia, a közgazdaságtan vagy a jogtudomány) meglátásaira is, azok módszertani eszköztárát és elméleti bázisát is bevonva a kutatás folyamatába.

fogalmi és esetelemzést, a fenomenológiát, a narrativizmust, a reflektív egyensúlyt, a specifikációt, az indokok mérlegelését, az etikai átvilágítást, az etikai menedzsmentet, az etikai mátrixot stb.” (saját fordítás)

Ennek tükrében alkottam meg a következő *hétosztatú modellt*, amely a dolgozat tudományos kontextusát meghatározza a továbbiakban, kihangsúlyozva az alkalmazott etika [1] gyakorlati jellegét; [2] interdiszciplináris természetét; [3] az elmélet és gyakorlat kölcsönhatását; [4] adaptivitását; [5] módszertani sokszínűségét; [6] jövőorientáltságát és [7] társadalmi szerepvállalását.



1. ábra. Az alkalmazott etika hétosztatú modellje (saját szerkesztés)

A modellhez kapcsolódó definíció szerint, az alkalmazott etika az alkalmazott filozófia azon ága, amely célzottan erkölcsi kérdésekkel, dilemmákkal és az *emberi cselekvés etikai aspektusaival foglalkozik* elemző módon, gyakorlati kontextusban és interdiszciplináris megközelítést alkalmazva. Az alkalmazott etika nem csupán az elméleti etika emberi praxisra vetített megvalósítását jelenti, hanem egy komplex, reciprok interakción alapuló metodológia, amely során az *elméleti megfontolások és a gyakorlati implementációk kölcsönösen alakítják egymást*. Ez a kölcsönhatás elősegíti, hogy az etikai teóriák a folyamatosan változó, valóságos kontextusokhoz igazodjanak, miközben a gyakorlati tapasztalatok újraértelmezik és finomítják az elméleti kereteket, ezáltal téve az alkalmazott etikát *dinamikussá és adaptívá* a változó társadalmi-technológiai környezetben. Az alkalmazott etika bár támaszkodik filozófiai elméletekre, pragmatikus megközelítést alkalmaz, hangsúlyt fektetve a kontextus-specifikus irányelvekre, esettanulmányokra és narratívákra, *folyamatos párbeszédet fenntartva a természet-, társadalom és humán tudományokkal*, sok esetben azoktól kölcsönzött változatos módszertani eszköztárt alkalmazva. Az alkalmazott etika *jövőorientált perspektívát kínál*, a jelen kihívásaira koncentrálva, de a technológiai és társadalmi fejlődés potenciális következményeit is figyelembe véve vizsgálja és elemzi tárgyát, azzal az igénnyel lépve fel,

hogy befolyásolja a döntéshozatali folyamatokat és iránymutatást adjon az etikus cselekvésre vonatkozóan különböző területeken, hozzájárulva ezzel egy etikusabb társadalom kialakításához.

2.3 Felelősség a nyugati gondolkodásban

A felelősség – kortárs, komplex értelmében – viszonylag új fogalom a filozófiában és a társadalomtudományokban. A felelősség koncepciójának evolúciója szorosan összefonódik a modern társadalmak fejlődésével és az emberi cselekvés hatókörének bővülésével. Az ipari forradalom, a globalizáció és a technológiai fejlődés olyan új kihívásokat teremtett, amelyek a felelősség fogalmának újraértelmezését és kiterjesztését követelték meg. A felelősség fogalmának komplexitását tovább árnyalja, hogy különböző diszciplínák eltérő módon közelítik meg és értelmezik azt: míg a jogtudomány elsősorban a cselekvések következményeire és a számonkérhetőségre koncentrálnak, addig a filozófia és az etika a morális dimenziókat, az egyéni és kollektív felelősségvállalás kérdéseit vizsgálja mélyrehatóbban. Definíciós keretei képlékenyek, és eltérő tartalmakat jelölhet, attól függően, hogy jogi terminológiában vagy kortárs környezetetikai kontextusban használják. A doktori tézisemben a felelősség fogalmát kifejezetten az erkölcsi felelősség értelmében használom. Az erkölcsi felelősség, ahogy e dolgozat fókuszában áll, nem csupán a múltbeli cselekedetek értékelésére vonatkozik, hanem magában foglalja a jövőbeli következmények előrelátásának és a proaktív felelősségvállalásnak a lehetőségét is, ami különösen releváns a vizsgálatunk közvetlen tárgyát képező alkalmazott – környezeti-, bio-, vagy mérnöki – etika kontextusában, ahol a hosszú távú következmények és a jövő generációk iránti felelősség központi szerepet játszik.

2.3.1 Szabadság, ágencia és felelősség

Arnold S. Kaufman a Thomson Gale Encyclopedia of Philosophy erkölcsi felelősséget tárgyaló szócikkében kiemeli, hogy amennyiben meg akarjuk határozni az erkölcsi felelősség megítéléséhez szükséges és elégséges feltételeket, egyúttal értelmet kell adunk magának a kifejezésnek. Az emberek többsége az erkölcsi felelősség megállapításakor alapvetően az alábbi keretekben gondolkodik: „*A person is regarded as morally responsible for some act or occurrence x if and only if he is believed (1) to have done x, or to have brought x about; and (2) to have done it or brought it about freely.*”⁶ (Kaufman, 2006, p. 445).

⁶ „Egy személy akkor és csak akkor tekinthető morálisan felelősnek valamely x cselekedetért vagy eseményért, ha úgy hisszük, hogy (1) ő követte el x-et, vagy ő idézte elő x-et; és (2) szabadon követte el vagy idézte elő azt..” (saját fordítás)

Számolnunk kell azonban azzal, hogy mit értünk emberi cselekvés alatt; mi számít egy eredmény elérésének; és mindenekelőtt, hogy milyen értelemben használjuk a szabad, szabadon, szabadság kifejezéseket – mert ezek a fogalmak nagyon eltérő felelősségelméletekhez vezetnek, és értelmezésük alapvetően meghatározza az erkölcsi felelősség fogalmának mibenlétét, mivel a szabadság jelentése szorosan összefügg az egyén morális álláspontjával és igazságosságról alkotott felfogásával.

William K. Frankena abból indul ki, hogy az emberek általában szabadon választhatnak cselekedeteik között, másrészt, hogy ezek a választások és cselekedetek észszerűen előrejelezhetők, és nem pusztán véletlenszerű események eredményei. A szerző szerint morálisan igazolható az emberek felelősségre vonása, dicsérete vagy elmarasztalása, amennyiben szabadon cselekednek abban az értelemben, hogy saját vágyaik és jellemük kifejezésekként követik el tetteiket, saját meggyőződéseik fényében, és tisztában vannak tetteik következményével (Frankena, 1973, pp. 75-78). A felelősség megállapításához ismernünk kell az egyén szándékait, a cselekvés kontextusát és az erkölcsi normákat, amelyeket alkalmazunk, mivel ezek mind hozzájárulnak ahhoz, hogy pontosan meghatározzuk a felelősség mértékét és jellegét. Gary Watson *Two Faces of Responsibility* című tanulmányában kiegészíti ezt azzal, hogy beemeli az erkölcsi felelősség két különböző perspektíváját, az önkifejezés (self-disclosure) és az elszámoltathatóság (accountability) nézőpontját. A felelősség önkifejezésként való felfogása arra az elképzelésre épül, hogy tetteink a saját belső értékeink és céljaink kifejezései, míg az elszámoltathatóság nézőpontja ezzel szemben a társadalmi szabályozás és igazságszolgáltatás keretein belül vizsgálja a felelősséget, és az egyének közötti köteleességek és elvárások érvényesítésére, valamint az ezek be nem tartása miatti szankciók kivetésére összpontosít. Watson szerint mindkét nézőpont fontos szerepet játszik az erkölcsi életben, azonban kiemeli, hogy erkölcsi felelősség fogalmának központi eleme az elszámoltathatóság. E megkülönböztethető perspektívák mögött két, az etikai élethez központi módon kapcsolódó érdekrendszer húzódik meg. Az egyik az azzal kapcsolatos érdekeinket foglalja magában, hogy miként élünk jó emberi életet, milyen modellek és ideálok szerint. A másik a társadalmi szabályozáshoz, valamint a retributív és kompenzációs igazságossághoz kapcsolódó érdekeket öleli fel, és bár e két érdekrendszer teljesen nem választható el egymástól, fontos belátni, hogy ezeknek eltérő forrásaik vannak. Ez a megközelítés lehetővé teszi az erkölcsi felelősség fogalmának árnyaltabb értelmezését, és rávilágít a mögötte húzódó komplex érdekrendszerekre és perspektívákra (Watson, 1996, pp. 242-243).

R.A Duff a Routledge Encyclopedia of Philosophy (REP) felelősség szócikkében úgy fogalmaz, hogy én vagyok felelős a cselekedeteim közvetlenül szándékolt eredményeiért azáltal, hogy a cselekedeteimet a megvalósításukra irányítom, illetve felelős vagyok mindazon eseményekért és eredményekért, amelyek a cselekedeteimből következnek. A felelősség két fő típusát különböztethetjük meg: a *prospektív felelősség* arra a felelősségre vonatkozik, amely már egy cselekvést megelőzően is adott, ez együtt jár a cselekvés előzetes kiértékelésével - mindazon dolgok, amelyekkel cselekvőként nekem foglalkoznom kell, vagy figyelembe kell vennem. Ilyen felelősség például a természeti környezet megóvása, a jövő nemzedék iránti kötelezettségek, illetve egyes szakmatikai kérdésekhez kapcsolódó felelősségkörök. A *retrospektív felelősség* az, amelyet a cselekvés után viselek, olyan eseményekért vagy eredményekért, amelyek nekem, mint ágensnek tulajdoníthatók: ez egy cselekvés következményeinek utólagos felmérésével jár együtt. Lényegében visszamenőlegesen felelős vagyok azért, amit teszek vagy elmulasztok tenni, amikor feladataimat teljesítem (Duff, 1998, p. 7436, Szilágyi, 2022, pp. 115-116).

Amitai Etzioni, amerikai szociológus az 1988-ban megjelent *Moral Dimension* című könyvében azt állítja, hogy az emberi motivációt egyszerre befolyásolja két egymásra nem visszavezethető tényező: a hasznosság és a morál. Az egyéni döntéshozatal felfogható egyfajta egyensúlykeresésként, vagyis döntéseinket egyszerre határozzák meg haszonelvű és etikai szempontok: az egyén igyekszik az önérdekét összhangba hozni a számára érvényes morális elvárásokkal. Egyes morális szempontok a döntéseket befolyásolhatják, mint belső erkölcsi parancsok, illetve mint egyfajta iránytűk, amelyek az egyes cselekvési lehetőségek erkölcsi tartalmának mérlegelését, és ezeknek az egyéni morális preferenciák tükrében történő választását segítik elő. Etzioni szerint a morális megfontolások és a haszonelvű mérlegelések hatással vannak egymásra, és az emberek is igyekeznek egyensúlyt találni köztük, vagyis az az egyén igyekszik az önérdekét egyensúlyba hozni az erkölcsi elvárásokkal (Etzioni, 1988). Kihangsúlyozza az egyéni jogok és a közösség iránti felelősség egyensúlyának fontosságát, a felelősséget közösségi keretben vizsgálja. Etzioni szerint a modern társadalmakban az egyének jogai és kötelezettségei közötti egyensúly megbomlott, és túl sok figyelem irányul az egyéni jogokra a közösségi felelősségek rovására „a társadalmi felelősségvállalás egyike azon alapvető erényeknek, amelyek megerősítésre szorulnak” (Etzioni, 1995, p. 11). A liberális kommunitarizmus általa kidolgozott koncepciója az egyéni jogok és a társadalmi felelősségvállalás között egy érzékeny, gondosan kialakított egyensúlyt kíván elérni, elmozdulva a társadalmi szintű *közjó* felé. Nem szabad szerinte elfeledkeznünk a közösségek

morális aspektusáról és arról sem, hogy az egyének jellemzően valamely csoport tagjaként viselkednek és nem önálló ágensként (Szilágyi, 2022, pp. 118-119).

2.3.2 Az erkölcsi felelősség fogalmának társadalmi és nyelvi vonatkozásai

Míg Etzioni az egyéni jogok és a társadalmi felelősség egyensúlyának fontosságát hangsúlyozta, Martin tovább árnyalja a felelősség fogalmát azáltal, hogy megvizsgálja annak különböző értelmezéseit és társadalmi kontextusait. Martin szerint a felelősség különböző helyzetekben és kontextusokban eltérően nyilvánul meg, és ezek a megnyilvánulások nem feltétlenül erkölcsi jellegűek és jelentőségűek, vagyis a felelősség fogalma nem feltétlenül tartalmaz erkölcsi dimenziót. Például a munkahelyi feladatok elvégzése vagy egy háztartási kötelesség teljesítése nem mindig jár együtt erkölcsi megítéléssel. Martin különbséget tesz a felelősség különböző értelmezései között, mint a kötelesség, az elszámoltathatóság és a cselekvés módja, ami hozzájárul a felelősség komplex fogalmának megértéséhez. Martin *Responsibility: A Philosophical Perspective* című tanulmánya az erkölcsi felelősség fogalmát elemzi különböző nézőpontokból, különös tekintettel annak társadalmi és nyelvi vonatkozásaira. Wittgensteint hívva segítségül rámutat arra, hogy a felelősség fogalma sokféleképpen használható és értelmezhető, attól függően, milyen kontextusban jelenik meg, ezért a felelősség definíciója és alkalmazása mindig egy adott helyzetben dől el, nem lehet egyetlen, végleges definíciót adni a fogalomra. Martin felveti, hogy a felelősség fogalmának tanulmányozása során érdemes az etnometodológiai megközelítést alkalmazni, amely azt vizsgálja, hogyan alakul ki a felelősség jelentése és megértése a mindennapi helyzetekben, a résztvevők közötti interakciók során. Ez a módszer a társadalmi cselekvéseket és azok hétköznapi megnyilvánulásait elemzi, ahelyett, hogy merev technikai definíciókra támaszkodna. Az etnometodológia szerint a felelősség értelmezése mindig az adott kontextusban, a konkrét szituációkban formálódik.

A tanulmány részletesen tárgyalja a felelősség fogalmának technikai és hétköznapi használatát is, különös tekintettel a jogi és szervezeti kontextusokra – ennek kapcsán rámutat, hogy míg a jogi eljárások és szervezeti szabályok gyakran rigid definíciókat használnak a felelősség meghatározására, addig a mindennapi életben a felelősség fogalma sokkal rugalmasabban és változatosabban jelenik meg. A technikai definíciók nem helyettesítik az egyéni értelmezési és döntési folyamatokat, amelyek mindig szükségesek a felelősség megállapításához. A tanulmányban a felelősség fogalmának négy alapvető értelmezését különbözteti meg:

- a) *A felelősség, mint kötelesség vagy kötelezettség*: valamilyen munka vagy feladat; a rám háruló feladatok összessége, mindazok a dolgok, amelyek elvégzéséért, illetve fenntartásáért felelős vagyok. A felelősség ilyen jellegű értelmezései sok esetben nagyon konkrétak, és az erkölcsi jelentőségük is eltérő lehet.
- b) *Felelősnek lenni valakiért és/vagy valamiért, illetve felelősséggel tartozni valaminek/vagy valakinek*. Ez a fogalom általánosabb értelmezése és az erkölcsi jelentősége itt is jelentős mértékben függ a felelősség tárgyától – pl. más következményekkel jár egy postai dolgozó napi munkája során a levelek kézbesítése, mint a rám bízott gyermekekről való gondoskodás.
- c) *Valaki vagy valami felelős valamiért, ami megtörtént*. Ez az értelmezés annak a mérlegelése, hogy mi *okozta* valaminek a bekövetkeztét, mi vezetett valamilyen esemény megtörténtéhez. Ez az értelmezés gyakran kerül elő olyan helyzetekben amikor valamilyen negatív esemény következik be vagy valamilyen probléma felmerül, de pozitív vagy semleges helyzetekben is adekvát értelmezés – lényegében egy *ok-okozati összefüggés* feltárása.
- d) *Felelősségteljesen cselekedni*, valamit felelősségteljesen tenni – kellő gondossággal és figyelemmel cselekedni. Ennél az értelmezésnél magát a felelősség fogalmát a cselekvés módjára vonatkozó határozószóként használjuk, hogy jelezzük a cselekvés módját (Martin, 2007, p. 27; Szilágyi, 2022, pp. 112-113).

Martin azt hangsúlyozza, hogy a felelősség fogalmának alapos megértéséhez figyelembe kell venni annak nyelvi és társadalmi kontextusait, és nem szabad kizárólag technikai definíciókra hagyatkozni. A felelősség mindig egy adott helyzetben nyeri el jelentését, és ezt a jelentést a résztvevők közötti interakciók alakítják ki.

2.3.3 Richard McKeon a felelősségről

McKeon 1956-ban publikált *The Development and the Significance of the Concept of Responsibility* című tanulmánya az Institut International de Philosophie két ülészaka közé beékelődött találkozó számára készült összefoglaló előadása a felelősség fogalmáról és egyben a téma egyik legrészletesebb tudományos tárgyalása (Guzzo, 1956).

Tanulmányának bevezetésében McKeon a felelősségnek három dimenzióját határozza meg: [1] van egy külső dimenzió, amely megjelenik a jogi és politikai kontextusban, amelyben az állam szankciókat szab ki az egyéni cselekedetekre, és amelyben a tisztviselőket és a kormányokat felelősségre vonják a politikai tetteikért; [2] létezik egy belső dimenzió, amely

etikai gondolkodásban jelenik meg, ahol az egyén figyelembe veszi a cselekedeteinek következményeit és azokat a kritériumokat, amelyek befolyásolják a döntéseit; [3] illetve van egy úgynevezett átfogó vagy reciprok dimenzió, amely a társadalom és a kultúra közegében mutatkozik meg, itt az értékek az egyén karakterének autonómiája és a kapcsolt civilizáció struktúrája szerint rendeződnek (McKeon, 1957, p. 5; Szilágyi, 2021, p. 40).

Rámutat, hogy a felelősség fogalma a filozófiában relatíve újkeletű, konkrét megjelenése a 18. századra tehető az angol, a német és a francia nyelvben egyaránt (*responsibility* 1787, *responsabilité* 1787, *Verantwortlichkeit* 1791), azonban a fogalom szinte „észrevétlen” marad a 19. századi vitákban. A filozófiatörténet korábbi korszakait vizsgálva, a felelősség kérdése szinte csak az emberi cselekvés elszámoltathatósága és az okozatiság kérdéskörére vonatkozik – állapítja meg, az ókorból példaként Platón és Arisztotelészt említve, kiemelve azt, hogy az igazságosság fogalma különösen az utóbbi filozófus esetében az etikát a politikához kapcsolja. A későbbi századokban a kötelesség (Kitioni Zénón, Szent Ágoston) vagy a hibáztathatóság (Aquinói Szent Tamás és Duns Scotus) fogalmi a meghatározók az etikában, majd Hobbes, Locke és Hume esetében a szükségszerűségeen alapul az erkölcs, és a normatív megkülönböztetések a számonkérhetőség külső eszközeiből (mint például a polgári és büntetőjog érvényesítése) eredeztethetők. McKeon a felelősség történeti kontextusának felvázolásában Pufendorf, Wolff és Kant munkásságára is utal, kiemelve a szabad cselekvés jelentőségét, amely természete meghatározza a cselekvések normatív megkülönböztetését, így általánosságban maga a cselekvés megítélése is elszakad a számonkérés és a büntetés külső eszközeitől (McKeon, 1957, pp. 13-17). A felelősség fogalmának első használatát John Stuart Mill-nek tulajdonítja, aki mivel el kívánt távolodni a szükségszerűségről, a szándékoktól és a szabadságról szóló vitáktól, egy a felelősségre vonás hagyományán alapuló jelentést adott neki, azzal érvelve, hogy a felelősség érzete a büntetés bekövetkeztének várható tényéből adódik: az egyén tudatosan vállalt rossz cselekvéseivel kívül helyezi magát a társadalmon, az emberek ellenszenvét kiváltva, ezáltal mindazon hasznoktól és jótéteményektől is elesik amelyet a közösség nyújtani képes számára - így a felelősségvállalás annak a tudata, hogy megérdemljük a büntetést (Szilágyi, 2023 pp. 387).

A 19. század végén Levy-Bruhl fenntartásokkal kezelte Mill megközelítését (Levy-Bruhl, 1884), ő egy észszerű szabad lény koncepciójára vezeti vissza a felelősséget, nála a felelősség a társadalomból ered, és a társadalom fejlődésével együtt fejlődött. Ebben a kontextusban meg kell különböztetnünk az objektív valósághoz kapcsolódó jogi felelősség fogalmát és a szubjektív és konkrét tartalomtól mentes erkölcsi felelősséget – ezt a gondolatot

idézi fel McKeon a bevezetőben, olyan elméletként, ami megtermékenyítően hatott a saját felelősség-koncepciójának kidolgozásában.

„*A politikai felelősség eszméje megelőzi a felelősség eszméjének kifejlődését*” (McKeon, 1957, p. 23) fogalmazza meg alapvetését McKeon, rámutatva arra, hogy a felelősség modern koncepciója a 18. század végi és a 19. század eleji etikai és politikai diskurzusban jelent meg először. McKeon szerint a „felelős kormányzat” eszménye egyszerre jelenik meg Nagy-Britanniában, Franciaországban és az Egyesült Államokban, és két alapvető aspektusa van: egyfelől a kormány felelős, ha olyan jogi keretek között működik, amelyben a hivatalos működés és annak ellenőrzése észszerűen kiszámítható, másfelől ha a kormányzás támogatja a nép azon elvárását, hogy olyan intézményeket tartson fenn és működtessen, amelyek biztosítják a személyes választás és a politikába való beleszólás lehetőségét (Szilágyi, 2021, pp. 39-40).

Fontos az a gondolat, hogy az egyén csak akkor lehet felelős amennyiben rendelkezik megfelelő politikai eszközökkel az adott kormány politikájának befolyásolására: „*a felelős kormányzás a felelős néptől függ*” – ez az egyéni felelősség kialakulásának első állomása, de a nép csak azáltal szerez felelősséget, hogy gyakorolja azt. A felelősség szabadság is egyben, amely a politikai közösségben fejlődött ki, az erkölcsi kritériumok szabad megválasztása a közösség erkölcsi érzékét és igazságosságát jobban előmozdítja és jobban segíti a racionális döntéseket szemben egy tekintélyelvű rendszerrel.

A politikai felelősség szoros kapcsolatot jelent a polgár és a politikai közösség között, és a politikai felelősség eszméje visszatükrözi az egyes társadalmak kulturális értékeit. A politikai felelősség általános gondolatát jelentősen befolyásolta a kulturális felelősség fogalma: a kulturális vagy társadalmi felelősségvállalás biztosítja a kapcsolatot, amely által a politikai és erkölcsi felelősség kölcsönösen befolyásolni képesek egymást (McKeon, 1957, p. 25). Az a felismerés, hogy a közösség felelősséggel tartozik más közösségek iránt, kiegészül azzal, hogy az egyén is elismeri tetteinek felelősségét és azt, hogy elszámoltatható. McKeon kiemeli, hogy az egyén felelőssége és annak a közösségnek a felelőssége, amelynek a tagja, interdependens viszonyban vannak. A korszak egyik nagy felismerése az, hogy az egyéni önérték hatékony érvényesítése káoszt eredményez, ha nincsenek jól definiált társadalmi elvek, szerepek vagy transzcendens ideálok, amelyeknek meg kell felelni (McKeon, 1957, p. 21; Szilágyi, 2021, p. 40).

McKeon hozzájárulása a felelősség fogalmának értelmezéséhez elsődlegesen a felelősség fogalmi eredetének feltárásában ragadható meg. Fontos felismerés, hogy az egyéni felelősség ma oly általánosnak számító erkölcsi formája nem előzte meg a kollektív

felelősséget. „Az erkölcsi felelősség kialakulása és elismerése előtt nem léteztek erkölcsös egyének” (McKeon, 1957, p. 28). A politikai és kulturális intézmények, valamint az egyén normái és döntései közötti kölcsönhatás az, ami a felelősség fogalmának változását munkálta

McKeon rámutat, hogy az erkölcsi felelősség a politikai és jogi felelősséggel szoros összefüggésében alakult ki: nem volt erkölcsi felelősség, amíg nem voltak közösségek, amelyekben az egyének felelősségre vonhatóvá váltak cselekedeteikért – ez a mozzanat, amikor az egyén morális cselekvővé vált. Az elszámoltathatóság és a beszámíthatóság jelöli ki azon kereteket, amelyek között az önálló ágens azonosíthatóvá válik az etikai rendszerekben. A jó és a rossz cselekedetek egyaránt felróhatóvá váltak a közösségben, amely egyúttal az elszámoltathatóság háttérét is biztosította. A felelősségrevonás és az elszámoltathatóság nehézségeire adott válaszainkkal, valamint az erkölcsi pluralizmus problémájának kezelésével új problémákat teremtettünk a felelősség tekintetében – ezeket a reflexív módon keletkezett problémákat úgy lehetne tisztázni, ha visszatérnénk a felelősség fogalmához és annak 19. század eleji eredetéhez (Szilágyi, 2023, pp. 388-389).

A felelősség fogalmának vizsgálata rávilágít annak összetett és változó természetére. A különböző elméleti megközelítések bemutatása jól szemlélteti, hogyan alakult a morális felelősség értelmezése a társadalmi, politikai és kulturális változásokkal párhuzamosan.

A dolgozatban ezek alapján a felelősséget egy komplex, dinamikus fogalomként határozom meg, amely az egyén cselekedeteinek és azok következményeinek erkölcsi és társadalmi értékelésére, valamint az egyén és a közösség közötti kölcsönös kötelezettségekre vonatkozik. Magában foglalja az egyén szabad döntéshozatalának képességét, a cselekedetek szándékoltságát és előreláthatóságát, valamint az elszámoltathatóságot mind egyéni, mind társadalmi szinten. A felelősség fogalma egyaránt tartalmaz előzetes mérlegelésre és kötelezettségvállalásra (prospektív) és utólagos számonkérhetőségre építő (retrospektív) elemeket. A felelősség koncepciója magában foglalja az egyéni autonómia és a közösségi normák közötti egyensúly keresését, és feltételezi az egyén képességét arra, hogy cselekedeteinek morális tartalmát mérlegelje és azokért vállalja a következményeket. A felelősség értelmezése szorosan összefügg a politikai, jogi és kulturális kontextussal, valamint az adott társadalom értékrendjével.

Ahogy McKeon elemzése is sugallja, a felelősség megértéséhez elengedhetetlen a történeti és kulturális kontextusok figyelembevétele, valamint az egyéni és kollektív dimenziók közötti kölcsönhatások vizsgálata. Ez a megközelítés lehetővé teszi számunkra, hogy a

felelősség fogalmát ne csak elméleti síkon értelmezzük, hanem annak gyakorlati alkalmazását is elősegítsük a következőkben tárgyalandó komplex etikai kihívások vizsgálatában.

2.4 Az innováció értelmezésének lehetőségei

A dolgozat közvetlen tárgyát képező innováció definícióinak fejlődése során a jelentés folyamatos bővüléséről beszélhetünk, amely szektorálisan kiegészül specifikumokkal, attól függően, hogy funkcionális vagy teoretikus kontextusban kerül használatra. Történeti (gazdaságtörténeti) kontextusban az innováció fogalmának definíciós kísérletei jól példázzák az értelmezés nehézségeit és tág korlátait, ezért különösen fontos, hogy a dolgozatban egy meghatározott jelentéstartalommal használjam ezt a fogalmat. Ez a rövid áttekintés azt szolgálja, hogy a tárgyról alkotott definíciós kísérletek számbavételével az általam használt definíció fő kereteit, és az értelmezés kritériumait is kijelöljem.

Benoit Godin – az egyik legismertebb innovációkutató, aki kutatásai révén jelentősen hozzájárult az innováció történeti fejlődésének és komplexitásának megértéséhez – az innováció mai értelmezésének történeti gyökereit vizsgálva rámutat, hogy a fogalom igazán az ipari forradalom generálta technikai fejlődéssel egy kontextusban kap figyelmet, majd használata a 20. században terjed el, és a gyökereinél az invenció és az imitáció fogalma is megjelenik (Godin, 2020a, pp. 5-6). Godin egy másik tanulmányában (2017, pp. 18-25) kifejti, ahogy az innováció fogalma a 19. század végén és a 20. század elején kezdett kialakulni, hamar központi szerepet kapott a társadalmi változások megértésében. Ezt megelőzően az innováció nem rendelkezett formális definícióval, és a fogalom csekély elméleti jelentőséggel bírt, leginkább nyelvi eszközként szolgált az ellenfelek kritizálására, azaz negatív konnotációval használták a kifejezést vitahelyzetekben. Arra hívja fel a figyelmet, hogy bár a 20. század előtt nem létezett az innováció tudományos igényű elmélete, a fogalom használata és megítélése már akkor is jelen volt a közbeszédben, elsősorban kritikai és pejoratív értelemben. Az innováció modern, pozitív értelmezése csak később, a 20. század folyamán alakult ki, de a korábbi negatív konnotációk még hosszú ideig hatással voltak az innovációról és az újítókról kialakult képre:

„The definition of what constitutes innovation and who qualifies as an innovator was set by those who opposed it: an innovator was seen as someone who advocates for deliberate change, in contrast to change ordained by God or dictated by political and religious authorities. Innovation was viewed as a threat to the established order. It was this negative portrayal of innovation that innovators—

whether revolutionaries, inventors, or others—had to challenge as they began to embrace the concept positively in the nineteenth century”⁷ (Godin, 2015, p. 46).

Godin az innováció két különböző szemantikai mezőjét azonosítja: az egyik a korábbi évszázadokra, míg a másik a huszadik századra vonatkozik. Az előző évszázadokban a szemantikai mező négy kulcsfogalomból állt: változás, újítás, reformáció és forradalom. Míg a változás általánosan elfogadott volt, addig az újítást és a változtatást szándékos beavatkozásnak tekintették, megkülönböztetve a természetes vagy Isten adta változásoktól. A huszadik században az innováció szemantikai mezője átalakult, megtartva néhány fogalmat az előző évszázadokból, mint például a változás és a forradalom, de új kifejezésekkel és értelmezésekkel is bővült. Az innováció az eredetiség szinonimájává vált, és az újítókra úgy tekintettek, mint az új gyakorlatok korai alkalmazóira, különösen azokra, akik elsőként hoztak forgalomba új találmányokat. A forradalmi vagy nagyszabású újítások eszméje a társadalomra gyakorolt vélt hatásuk miatt került előtérbe és kapott pozitív értelmezést (Godin, 2015, pp. 67-68).

Az innováció elmélete kezdetben az eredetiséget és az elsőbbséget hangsúlyozta, szembeállítva azt az utánzással (imitation). Az innováció értelmezésében sokáig domináns nézet szerint az innováció az eredeti ötletek vagy találmányok első kereskedelmi bevezetését jelenti – ez a megközelítés különösen jellemző volt a gazdasági és menedzsment elméletekre. Bár az innováció és az utánzás megkülönböztetése sokáig domináns maradt, ez a nézet egyre inkább megkérdőjeleződik, az újabb megközelítések szerint az utánzás és az adaptáció is lehet az innováció egy formája, nem csak az teljesen eredeti ötletek létrehozása (Godin, 2020a, pp. 5-6).

A neves kultúrantropológus Homer Barnett az előzőekhez kapcsolódó, az innovációval és kulturális változással kapcsolatos elméletét részletesen az *Invention and Cultural Change* című tanulmányában fejti ki. Barnett szerint az innovációk gyakran egy meglévő forma kölcsönzésére és annak egy új kontextusban való alkalmazására épülnek, az új formák általában az eredeti formák származékai, amelyeket az új környezet igényeihez igazítanak. Az innovációk tehát gyakran a régi formák kissé módosított változatai, amelyek idővel egyre jobban alkalmazkodnak az új funkciókhoz. Az innovációk elfogadása és terjedése összetett folyamat, amely számos tényezőtől függ, beleértve a közösség nyitottságát, az újítások hasznosságát és

⁷ „Annak meghatározását, hogy mi az innováció, és ki minősül újítónak, az ellenzők határozták meg: az újítót olyan személynek tekintették, aki szándékos változás mellett száll síkra, szemben az Isten által elrendelt vagy a politikai és vallási hatóságok által diktált változásokkal. Az innovációra úgy tekintettek, mint a fennálló rendet fenyegető veszélyre. Az újítónak - legyenek forradalmárok, feltalálók vagy mások - az innovációnak ezt a negatív képét kellett megkérdőjelezniük, amikor a XIX. században elkezdték pozitívan felfogni a fogalmat” (saját fordítás)

az újítók hitelességét (Barnett, 1942). Az 1953-ban megjelent *Innovation: the basis of cultural change* című nagy hatású művében, az innovációt a következőképp definiálja:

„An innovation is here defined as any thought, behavior, or thing that is new because it is qualitatively different from existing forms. Strictly speaking, every innovation is an idea, or a constellation of ideas; but some innovations by their nature must remain mental organizations only, whereas others may be given overt and tangible expression. "Innovation" is therefore a comprehensive term covering all kinds of mental constructs, whether they can be given sensible representation or not.”⁸ (Barnett, 1957, p. 7).

Barnett az innovációt minőségi különbségként határozza meg a meglévő formákhoz képest, hangsúlyozva, hogy az újítás nem csupán mennyiségi vagy fokozatos változást jelent, hanem alapvető minőségi eltérést a korábbi állapotokhoz képest. Kiemeli továbbá, hogy az innovációk nem korlátozódnak a fizikai tárgyakra vagy technológiákra, hanem magukban foglalják a gondolatokat, viselkedésformákat és egyéb szellemi konstrukciókat is. Ez a definíció kiszélesíti az innováció fogalmát, beleértve minden új gondolatot és ötletet, nem csupán azokat, amelyek gyakorlati dimenzióval bírnak.

Az 1930-as években a témakör első nagy teoretikusa Joseph Schumpeter elsődlegesen a gazdaság felől értelmezve a fogalmat, új termékek vagy eljárások bevezetését, új piacok és új szervezeti keretek létrehozását értette innováció alatt (Schumpeter, 1930), megkülönböztetve az innováció öt típusát: (1) új termelési folyamatok bevezetése, (2) új, a fogyasztók körében még nem ismert termékek vagy ismert termékek új minőségben történő előállítás, (3) új anyagok vagy erőforrások feltárása és bevezetése, (4) új piacok megnyitása, valamint (5) új szervezeti formák kialakítása (Schumpeter 1934, p. 66). Hangsúlyozta, hogy az innováció nem csupán technológiai újításokat jelent, hanem magában foglalja az üzleti modellek, piaci stratégiák és szervezeti formák kialakítását és megújítását is, amelyek mind hozzájárulnak a dinamikus gazdasági fejlődéshez. Schumpeter a kapitalizmus lényegét nem a statikus egyensúlyban, hanem a folyamatos változásban és innovációban látja: az innováció és a *kreatív rombolás* a rendszer alapvető jellemzői, nem pedig külső zavaró tényezők. A kapitalista rendszer fő hajtóereje nem külső tényezőkben rejlik, hanem belső impulzusokban: új termékek, technológiák, piacok és szervezeti formák megjelenésében. Ezek a folyamatos újítások alapjaiban alakítják át a gazdaság szerkezetét, lebontva a régi struktúrákat és újakat hozva létre

⁸ „Az innovációt úgy határozzuk meg, mint bármely gondolat, viselkedés vagy dolog, amely azért új, mert minőségileg különbözik a már létező formáktól. Szigorúan véve minden innováció egy ötlet vagy ötletek együttese; azonban egyes innovációk természetüknél fogva csak szellemi alakzatok maradnak, míg mások kézzelfogható, látható kifejezést nyerhetnek. Az 'innováció' ezért egy átfogó kifejezés, amely minden fajta szellemi konstrukciót magában foglal, függetlenül attól, hogy azok érzékelhető formában megjeleníthetők-e vagy sem.” (saját fordítás)

helyettük, és minden vállalkozási stratégiát e folyamatos változás kontextusában kell értelmezni, nem pedig egy állandó, változatlan környezet feltételezésével. Schumpeter szerint a valódi verseny nem az árakról szól, hanem az új termékekről, technológiákról és szervezeti formákról, amelyek alapjaiban rengetik meg a meglévő vállalatok pozícióját. Ez a fajta verseny nem csak akkor hat, amikor ténylegesen megjelenik, hanem már a fenyegetettség érzése is ösztönző erővel bír a vállalkozásokra nézve (Schumpeter, 1942, pp. 82-85)

Az innováció fogalma később is jellemzően megmarad a gazdaságtudományok diszciplináris keretei között és elsődlegesen új eszközök, termékek, szolgáltatások, felhasználási lehetőségek és működési módok feltárását, implementációját és adaptációját értik alatta a témával foglalkozó kutatók (Downs & Mohr, 1976; Drucker, 1958; Howard & Sheth, 1969). Seaden és Manseau (2001) dinamikus tevékenységként tekintenek az innovációra, amely az ő értelmezésükben az a kreatív folyamat, amelynek során új vagy továbbfejlesztett ötletek sikeres kidolgozását követően azok alkalmazása új eredményt hoz létre, amelyek a gyakorlatiasak (hasznosíthatók) és értéket képviselnek – vagyis az ötletek valamiféle értéké alakulnak át.

Az innováció legismertebb funkcionális definíciója az OECD által közreadott ún. Oslo Kézikönyv 2005-ös verziójából származik, amely egy nemzetközi iránymutatást/javaslatot ad az innováció mibenlétének meghatározásához, miszerint az *„innováció új vagy jelentősen továbbfejlesztett termék (áru/szolgáltatás) vagy eljárási módszer/gyakorlat bevezetése”*. A 2018-ban közreadott negyedik kiadásában egy kibővített, meglehetősen technokrata definícióval találkozunk: *„a new or improved product or process (or combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process)”*⁹ (OECD–European Commission, 2008, p. 22).

The Oxford Handbook of Philosophy of Technology az innovációt olyan tevékenységek összességéként definiálja, amelyek új vagy jelentősen továbbfejlesztett termékeket vagy folyamatokat eredményeznek. Az innováció lehet technológiai jellegű, de vonatkozhat hagyományos termékekre, szolgáltatásokra, marketing módszerekre, kereskedelmi gyakorlatokra, eljárásokra vagy szervezeti formákra is. A technológiai innováció új koncepciók, technikák és tervek feltalálását, ezek termékekké alakítását, piacra vitelét, majd a társadalmi és

⁹ *„Az innováció egy új vagy továbbfejlesztett termék vagy eljárás (vagy ezek kombinációja), amely jelentősen különbözik az innovációt megvalósító szervezet által a potenciális felhasználók számára elérhetővé tett vagy általa használt korábbi termékektől vagy eljárásoktól”* (saját fordítás)

gazdasági gyakorlatba való beépítését jelenti, és túlmutat a pusztán feltaláláson, mivel magában foglalja a megvalósítást, a termékfejlesztést, a marketinget és a társadalmi terjesztést is. A szócikk szerzője Philip Brey megjegyzi, hogy nem minden technológiai tervezés innovatív, a rutin tervezés jól definiált keretek között zajlik, meglévő prototípusok alapján, és nem igényel jelentős innovációt vagy kreativitást. Ezzel szemben az innovatív és kreatív tervezés lényegesen új terveket vagy megoldási elveket hoz létre, és e kettő között helyezkednek el az újratervezés különböző formái, amelyek során egy meglévő tervet fejlesztenek tovább új követelmények teljesítése vagy a teljesítmény javítása érdekében (Brey, 2021, pp. 400-401).

A már többször említett Godin különbséget tesz az innováció két modellje között: az egyik a *lineáris modell*, amely szerint a technológiai innováció az alapkutatástól az alkalmazott kutatáson át a fejlesztési fázisig tart; a másik a *szekvenciális modell*, amely az innovációt egy időben zajló folyamatként értelmezi a feltalálástól (invention) az elterjedésig (diffusion). A szekvenciális modell gyökerei a 19. századi antropológiai vitákra vezethetők vissza, ebben az időszakban alakult ki egy ellentét az invenció (vagyis egy csoport által kifejlesztett) és a diffúziós (átterjedő, más kultúrákból átvett) kultúraváltozás-elméletek között. E vita feloldásaként született meg az a gondolat, hogy a feltalálás és az elterjedés egyazon folyamat egymást követő lépései. Ez a megközelítés jelentős hatással volt az innováció későbbi értelmezésére, és az 1920-as évektől kezdve az antropológusok és szociológusok írásaiban egyre gyakrabban jelent meg az innováció szekvenciális folyamatként való értelmezése – bár ahogy Godin rámutat – a kutatók gyakran előnyben részesítették az egyik vagy a másik aspektust. Megjegyezi, hogy napjainkban ismét változás figyelhető meg: az innováció fogalma egyre inkább a találmányok diffúzióját és kereskedelmi hasznosítását hangsúlyozza, míg a feltalálás szerepét kevésbé tartják jelentősnek (Godin, 2014b).

Lundvall hangsúlyozza az innováció mindenütt jelenlévő és kumulatív jellegét, és a megközelítésében elmosódik a határvonal a feltalálás (invention), az innováció és a diffúzió között. Az innováció nem egyetlen eseményként, hanem folyamatként jelenik meg, amelynek időbeli behatárolása és pontos meghatározása nehézkes. Úgy látja, hogy a modern kapitalizmusban az innováció alapvető és állandó jelenség, amely nélkülözhetetlen a vállalatok és nemzetgazdaságok hosszú távú versenyképességéhez. Az innováció a modern gazdaság minden területén jelen van, a folyamatos tanulás, kutatás és felfedezés formájában, ami új termékeket, technológiákat, szervezeti formákat és piacokat eredményez, még ha néhol lassú és fokozatos is a folyamat. Az innováció gyakran meglévő lehetőségek és komponensek új kombinációjaként értelmezhető: szinte minden innováció már létező tudást ötvöz új módon, bár

az új kombináció megtalálása olykor hatalmas szellemi erőfeszítést vagy kivételes kreativitást igényel. Az innováció néha radikális törést jelent a múlttal, elavulttá téve a felhalmozott tudás egy részét. Ez a „kreatív rombolás” folyamata, amely nemcsak a termelés, de a tudás szerkezetét is érintheti (Lundvall, 1992, pp. 8-9).

A szociológus Everett Rogers, az innováció fontos elméletalkotója szerint „(...) *innovation, defined as an idea, practice, or object that is perceived as new by an individual or another unit of adoption. An innovation presents an individual or an organization with a new alternative or alternatives, with new means of solving problems*”¹⁰ (Rogers, 1983, p. XVIII-XIX). Rogers öt kulcsfontosságú jellemzőt azonosít, amelyek hatással vannak az új ötletek és technológiák elterjedésére, ezek a relatív előny (relative advantage), a kompatibilitás (compatibility), az összetettség (complexity), a kipróbálhatóság (trialability) és a megfigyelhetőség (observability). Szintetizálva ezt az öt jellemzőt, azt sugallja, hogy azokat az innovációkat, amelyeket előnyösebbnek, kulturálisan összeegyeztethetőbbnek, egyszerűbben használhatónak, könnyen tesztelhetőnek és láthatóan hatásosnak érzékelnek, általában gyorsabban fogadják el, gyorsabban terjednek el – emellett elismeri, hogy bár ezek jelentős tényezők, nem az egyetlenek, amelyek befolyásolják az elfogadási arányokat (Rogers, 1983, pp. 15-17).

Jack Andrew Morton az *The Innovation of Innovation* című tanulmányában az innovációt összetett, integrált folyamatként, nem elszigetelt tevékenységek sorozataként mutatja be. Az innováció a tudományos felfedezéstől a piac megteremtéséig több, egymással összefüggő szakaszt foglal magában, amelyek mindegyike összehangoltan dolgozik egy közös cél érdekében, vagyis az innováció nem egy egyszerű, lineáris folyamat, az innováció nem csak az egyik elemet jelenti, hanem mindezek szinergikusan együttműködő integrációját. Ez a szemlélet megkérdőjelezi az innovációról alkotott leegyszerűsítő nézeteket, és kiemeli annak sokrétű jellegét. A szerző úgy tárgyalja az innováció fogalmát, mint „céltudatos emberi folyamatot”, hangsúlyozva az emberi tényezőt a technológiai fejlődésben: „*All the parts and couplings of this process stem from the individually creative, yet cooperative actions of people. I like to call it a »people process with a purpose«*”¹¹ (Morton, 1968, p. 57) – ezzel kiemeli az egyéni kreativitás és az együttműködés fontosságát az innováció előmozdításában, azt sugallva,

¹⁰ „Az innováció olyan ötlet, gyakorlat vagy tárgy, amelyet egy egyén vagy egy másik befogadó egység újként érzékel. Az innováció új alternatívát vagy alternatívákat kínál az egyén vagy szervezet számára, új eszközöket biztosítva a problémák megoldására.” (saját fordítás)

¹¹ „E folyamat minden része és kapcsolódási pontja az emberek egyénileg kreatív, de együttműködő tevékenységéből ered. Szeretem ezt »céltudatos emberi folyamatnak« nevezni.” (saját fordítás)

hogy bár a technológiai innováció tudományos és technikai információkkal foglalkozik, alapvetően az emberi leleményesség és együttműködés hajtja. Morton az innovációt tudásintenzív törekvésnéként írja le, hangsúlyozva az információkezelés és a tudás létrehozásának kritikus szerepét a technológiai fejlődés előmozdításában. A technológiai innovációt a következőképp definiálja: „(...) *perception, creation, and transformation of relevant science into new and improved products and services*”¹² (Morton, 1968, p. 57), kiemelve a tudományos ismeretek gyakorlati, piacképes innovációvá alakításának folyamatát.

Csizmadia Tibor hasonlóan definiálja az innovációt, amit az Innováció versus minőségmenedzsment – avagy az oximoron feloldása címet viselő könyvében fogalmaz meg:

„egy olyan gazdasági vagy társadalmi relevanciával bíró újdonság létrehozása, átvétele vagy felhasználása, amely relatív hozzáadott értékkel bír; vonatkozhat termékek, szolgáltatások, folyamatok, piacok vagy iparágak megújítására vagy kibővítésére, új eljárások és módszerek kifejlesztésére, illetve új menedzsmentrendszerek kialakítására. Tehát az innováció egyszerre folyamat és végeredmény” (Csizmadia, 2023, s.p.).

A Handbook on Alternative Theories of Innovation szerkesztői bevezetője rámutat, hogy az innováció az elmúlt évtizedekben kiemelt fontosságú témává vált a tudományt művelők és a döntéshozók számára, és ezzel párhuzamosan az innováció fogalma is jelentős átalakuláson ment keresztül. Az innovációval kapcsolatos elméletek több tudományterületen párhuzamosan fejlődtek, és töltődtek meg az adott diszciplína sajátosságainak megfelelő tartalommal, azonban ezen hagyományosan közgazdasági innováció-elméletek mellett mára megjelentek alternatív megközelítések is, amelyek a mainstream technológiai-ipari innovációfelfogást kiegészítik vagy egyenesen megkérdőjelezzik, ilyenekre példák a társadalmi innováció, a felelős innováció vagy a fenntartható innováció koncepciói. Ezen új megközelítések azonban sok esetben régebbi koncepciók újracsomagolásai, emiatt egy kritikai perspektívából kell vizsgálnunk és kezelnünk őket – hívják fel a figyelmet (Godin et al., 2021, pp. 1-9). Az innováció fogalmának összetettségét jól mutatja be az általuk megadott többértelmű értelmezés:

„What do we mean by ‘innovation’? Innovation has two usages. To the moderns, innovation is novelty, novelty of any kind: new ideas, things, behaviors and practices. In this sense, innovation is a synonym for novelty and invention. But novelty is only one of the meanings of innovation. As the etymology suggest (in+), innovation is the introduction of something new into practice. To many, innovation excludes some types of novelty: the mental or speculative. Innovation is contrasted to contemplation. Innovation is action. What distinguishes innovation from change

¹² „a releváns tudományos ismeretek érzékelése, létrehozása és átalakítása új és továbbfejlesztett termékekké és szolgáltatásokká.” (saját fordítás)

*is purpose, to mainstream theorists at the least. Innovation is designed change, a deliberate initiative. A discovery or an invention becomes an innovation only when it is introduced, applied, adopted, distributed, maybe commercialized, that is, only if it is used and useful*¹³ (Godin et al., 2021, p. 3).

Bessant és Phelps is rámutatnak, hogy az innováció egy komplex és többdimenziós jelenség, amelynek vizsgálata és elemzése interdiszciplináris megközelítést igényel. Emiatt kihívást jelent a különböző nézőpontok és módszerek szintetizálása egy átfogó keretrendszerbe. A szerzők kiemelik, hogy ez a sokféleség különösen kihívást jelent az innovációs irodalom szintetizált keretrendszerbe történő összefoglalásakor, mivel az egyes kutatók által alkalmazott megközelítések jelentősen eltérhetnek egymástól, és a széleskörű téma kezelése komplexitást von maga után (Bessant & Phelps, 2006, p. 22).

Juett R. Cooper egy multidimenzionális megközelítés alkalmazását javasolja az innováció jobb megértéséhez és menedzseléshez. Meglátása szerint az innováció fogalma azért problematikus, mivel túl általánosan használják, egydimenziós kategóriákba sorolják és a definíciók széles skálán mozognak, a legszűkebb értelmezéstől, amely az innovációt a találmánnyal azonosítja, egészen a legtágabb megközelítésig, amely szerint bármilyen, az alkalmazó által újnak tekintett ötlet, gyakorlat vagy tárgy innovációnak számít. A nem megfelelő definíciók megnehezítik az összehasonlítást, a pontos meghatározás hiánya könnyen vezet ellentmondásos és nem egyértelmű következtetésekhez az innováció vizsgálata és kezelése során. Éppen ezért kulcsfontosságú, hogy minden egyes kutatás vagy gyakorlati alkalmazás esetén pontosan definiáljuk, mit értünk innováció alatt, – ehhez a megoldást egy közös viszonyítási alap és elméleti keret megléte szolgáltatná. Cooper végkövetkeztetése, hogy az innováció jobb megértéséhez és menedzseléséhez egy multidimenzionális megközelítésre van szükség. Az innováció különböző dimenzióinak együttes figyelembevétele segíthet feloldani a korábbi kutatások és elméletek ellentmondásait, és pontosabb képet adhat az innováció és a szervezeti jellemzők kapcsolatáról, mivel a legtöbb innováció valóban több dimenzió mentén értelmezhető egyszerre. A tanulmányban javasolt modell legalább három fő dichotómia mentén vizsgálja az innovációkat: termék vs. folyamat, radikális vs. inkrementális,

¹³ „Mit értünk innováció alatt? Az innováció kétféleképpen használható. A modernnek számára az innováció az újdonság -bármilyen újdonság: új ötletek, dolgok, viselkedésmódok és gyakorlatok. Ebben az értelemben az innováció az újdonság és a találmány szinonimája. Az újdonság azonban csak az innováció egyik jelentése. Ahogy az etimológia is sugallja (in+), az innováció valami újnak a gyakorlatba való bevezetése. Sokak számára az innovációt kizárja az újdonság bizonyos típusait: a szellemi vagy spekulatív újdonságot. Az innovációt szembeállítják a szemlélődéssel. Az innováció cselekvés. Ami megkülönbözteti az innovációt a változástól, az a cél, legalábbis a mainstream teoretikusok szerint. Az innováció tervezett változás, szándékos kezdeményezés. Egy felfedezés vagy találmány csak akkor válik innovációvá, ha bevezetik, alkalmazzák, elfogadják, elterjesztik, esetleg kereskedelmi forgalomba hozzák, vagyis csak akkor, ha használják és hasznosítják.” (saját fordítás)

valamint technológiai vs. adminisztratív. Ez a modell lehetővé teszi az innovációk többdimenziós térben való értelmezését, elismerve, hogy egy adott innováció egyszerre több dimenzió jellemzőit is hordozhatja (Cooper, 1998).

Vincent Blok 2021-es tanulmányában egy átfogó filozófiai megközelítést mutat be az innováció jelenségének komplexebb megértéséhez, kihívást intézve a hagyományos technológiai-gazdasági értelmezésekkel szemben. A szerző négy fő dimenzióban vizsgálja az innovációt, összekapcsolva annak folyamatát és eredményét, valamint ontikus és ontológiai szintjeit. Blok értelmezésében az innováció eredménye kettős természetű: ontikus szinten új termékeket vagy szolgáltatásokat jelent (például streaming szolgáltatások), míg ontológiai szinten egy teljesen új politikai-gazdasági világrendet hoz létre (például a digitális hálózatok világa). Ez a megkülönböztetés segít megérteni, hogy az igazi radikális innováció nem csupán új termékeket, hanem új létmódokat is teremt (hangsúlyos ontológiai dimenziója van). Az innovációs folyamat ontogenetikus természetű, ami azt jelenti, hogy nem lineáris vagy előre tervezhető, hanem iteratív és emergens. Ez a folyamat magában foglalja mind a kreativitást (negentrópia), mind a rombolást (entrópia), és nem tulajdonítható kizárólag egyéni szereplőknek vagy bemeneti tényezőknek, és az innováció újdonsága sem elsősorban az egyedi termékekben vagy szolgáltatásokban rejlik, hanem abban az „új világrendben”, amelyet létrehoz. Ez az újdonság nem a semmiből jön létre, hanem a történelmi folyamatokba ágyazódva jelenik meg, egyszerre mutatva folytonosságot és szakítást a múlttal. Az emberi részvétel az innovációban átértékelődik, és az ember nem elsődleges alkotóként, hanem egyfajta együtt-teremtő képességgel rendelkező résztvevőként jelenik meg. Ez a koncepció túlmutat az egyéni kreativitás hagyományos felfogásán, és az embert a világépítő folyamat részeseként értelmezi: *„Human creativity is involved in this ontogenetic process as co-creative capacity, namely human identity as an ontological outcome of innovation and at the same time human identity as pre-individual involvement in the process of innovation”*¹⁴(Blok, 2021, pp. 365-366). A szerző hangsúlyozza, hogy ez a megközelítés lehetővé teszi a kutatók számára, hogy túllépjenek – az előzőekben tárgyalt – hagyományos lineáris és körkörös innovációs modelleken, és segíthet jobban megérteni az olyan komplex jelenségeket, mint a társadalmi-, a fenntartható- és a felelősségteljes innováció kérdése. Blok rámutat, hogy ez a filozófiai megközelítés új utakat nyithat az innováció etikájának vizsgálatában is, túllépve a pusztán technológiai és gazdasági megfontolásokon, az innovációt, mint emberi ethoszt vizsgálva.

¹⁴ „Az emberi kreativitás ebben az ontogenetikus folyamatban együtt-teremtő képességként van jelen, nevezetesen az emberi identitás, mint az innováció ontológiai eredménye, és ugyanakkor az emberi identitás, mint pre-individuális részvétel az innovációs folyamatban.” (saját fordítás)

Az innováció kortárs értelmezéseit áttekintve, én jelen dolgozatban a következő saját definíciót tekintem irányadónak az innováció mibenlétére vonatkozóan, és az innovációs tevékenységeket ennek a kontextusában tárgyalom:

Az innováció egy *multidimenzionális folyamat*, amely magában foglalja új vagy jelentősen továbbfejlesztett elméletek, termékek, szolgáltatások, technológiák, folyamatok, szervezeti formák vagy piaci megközelítések *létrehozását, adaptációját, implementációját és diffúzióját*. Ez a folyamat lehet *fokozatos (inkrementális)* vagy *radikális*, és kiterjedhet *tudományos, technológiai, adminisztratív, társadalmi vagy gazdasági területekre*. Az innováció *minőségileg különbözik a meglévő formáktól*, relatív *hozzáadott értékkel* bír, és potenciálisan megfigyelhető, érthető változást eredményez az adott területen. Az innováció nem korlátozódik az eredeti ötletre vagy találmányokra, hanem magában foglalja azok *sikeres alkalmazását*, valamint a meglévő elemek *új kombinációit és adaptációit* is, amelyek értéket teremtenek és/vagy problémákat oldanak meg egy adott kontextusban. Az innováció továbbá ontogenetikus természetű, nem csupán új termékeket vagy szolgáltatásokat eredményez, hanem mélyreható változásokat hozhat létre a társadalmi-gazdasági rendszerekben, amely során új létezési módok és rendszerek jönnek létre, gyakran előre nem látható módon, a meglévő struktúrákkal kölcsönhatásban.

Ez a definíció átfogóan ragadja meg az innováció lényegét, integrálva a különböző elméleti megközelítéseket, figyelembe véve többek között Godin, Barnett, Blok és Schumpeter munkásságát. A definíció egyaránt hangsúlyozza az innováció folyamat- és eredményjellegét, valamint kiemeli a relatív hozzáadott érték és a kontextus fontosságát, továbbá inkluzív jellege révén nem csupán az eredetiséget, hanem az adaptációt és utánpótlást is magában foglalja. Multidimenzionális természete lehetővé teszi, hogy lefedje az innováció különböző típusait és formáit, így alkalmas a termék-, folyamat- és szervezeti innovációk besorolására is, emellett a rugalmassága és átfogó jellege biztosítja, hogy különböző szektorokban és kutatási területeken is alkalmazható legyen.

2.5 Összefoglalás

Az alkalmazott filozófia és etika fontos szerepet játszik korunk összetett problémáinak megértésében és kezelésében. E területek rugalmas és interdiszciplináris megközelítése lehetővé teszi, hogy a hagyományos filozófiai fogalmakat és módszereket gyakorlati kérdések vizsgálatára alkalmazzuk. Az alkalmazott etika közvetlenül foglalkozik az emberi cselekvés erkölcsi vonatkozásaival a mindennapi élet és az egyes szakmaterületek gyakorlatának a

kontextusában. A technológiai fejlődés gyors üteme és a modern társadalmak mindenki számára megtapasztalhatóan növekvő összetettsége, számos olyan etikai kihívást eredményez, amelyek az erkölcsi érvelés és bizonyos esetekben a döntéshozatal új megközelítését igénylik. Az alkalmazott filozófia és az alkalmazott etika hasznos gyakorlati eszközként jelennek meg a kortárs erkölcsi kérdések terén való eligazodáshoz. Az absztrakt filozófiai fogalmak és a valós problémák közötti szakadék áthidalásával az alkalmazott etika arra törekszik, hogy gyakorlatias útmutatást nyújtson a bonyolult háttérű döntésekkel és az egymással versengő értékekkel szembesülő egyének, szervezetek és politikai döntéshozók számára.

A felelősség koncepciója kiemelt jelentőségű mind az alkalmazott filozófiában, mind az etikában. Ez a fogalom magában foglalja az egyén szabad akaratát, döntéshozatali képességét, valamint tetteinek szándékoltságát és előreláthatóságát. A felelősség értelmezése azonban nem egységes, a különböző filozófiai irányzatok eltérő megközelítéseket kínálnak., és ez a sokszínűség tükrözi a fogalom összetettségét továbbá azt, hogy mennyire függ az értelmezés az adott kontextustól és nézőponttól. A felelősség már nem korlátozódik az egyéni elszámoltathatóság szűk területére, hanem a globalizáció és a konnektivitás korában egy sokkal tágabb, rendszerszintű értelmezést nyer. Ma már a felelősség fogalma nemcsak az egyén tetteinek közvetlen következményeit foglalja magában, hanem a társadalomra, a környezetre és a jövő generációira gyakorolt közvetett és hosszú távú hatásokat is, kiegészülve a kollektív felelősség komplex dimenzióival. A felelősségnek ez a kibővített értelmezése az etikai mérlegelés egy holisztikusabb megközelítését igényli, amely figyelembe veszi a kapcsolatok és interdependenciák összetett hálóját, amely a társadalmainkat jellemzi.

Az innováció filozófiai és kultúrtörténeti vizsgálata, amely a következő fejezet fókuszában áll, a fenti diskurzus organikus folytatásának tekinthető. Az innováció fogalma messze túlmutat a technológiai vagy gazdasági dimenziókon, gyökerei mélyre nyúlnak a nyugati bölcsélet talajába. Az innováció filozófiai perspektívából történő vizsgálata rávilágít arra, hogy az újítás képessége az emberi létezés egyik alapvető sajátossága, amely szorosan összefonódik olyan fundamentális kérdésekkel, mint az értékek, a haladás és a tökéletesedés eszméje. Az eszmetörténeti háttér vizsgálata hozzájárul ahhoz, hogy megértsük azokat a motivációkat, amelyek az emberiséget a történelem során előre vitték, legyen szó akár a tudományos felfedezésekről, a technológiai újításokról, a művészeti áramlatokról vagy a társadalmi reformokról; és egyúttal rávilágít arra is, hogy az innováció nem értéksemleges fogalom, hanem normatív implikációkkal bír, amelyek etikai reflexiót igényelnek.

III. Az innováció filozófiatörténeti kontextusa

A következő fejezetben az innováció eszmetörténeti kontextusba ágyazott vizsgálata során, szó esik majd arról, hogyan változott a fogalom értelmezése az évszázadok alatt, milyen filozófiai irányzatok befolyásolták az innovatív és az innovációról való gondolkodást és milyen etikai kérdések merülnek fel az újítások kapcsán. Ez a vizsgálat nem csak történeti szempontból érdekes, hanem segíthet megérteni jelenkori innovációs folyamatainkat és azok társadalmi hatásait is.

Miközben az üzleti világtól az akadémia közegéig egyforma súllyal nehezedik mindenkire az innovációs kényszer, az „innovációs kultúra” szervezeti kereteinek kiépítése és innováció kulcsfogalomként való válása a napi munkaszervezésben, a kultúránk újhoz és újításhoz való viszonyának vizsgálatára – néhány kivételtől eltekintve (Godin, 2012) – a filozófia kevés figyelmet szentelt. Az előzőekben láthattuk azt, hogy könyvtárnyi irodalom született az innováció tárgykörében, azonban a definíciós keretek bővülése és új kategóriák bevezetése csak látszólag visz közelebb minket a fogalom lényegének megragadásához – pedig amennyiben egy több mint két és fél ezer éves nyugati filozófiatörténeti idővonalon kívánjuk megtalálni az innováció gondolati csíráit, pontosabb irányokra lesz szükségünk.

Jelen fejezetben ennek a történeti ívnek a felvázolására teszek kísérletet, bemutatva a fogalom megjelenési formáit, kontextusait, kiemelve a bölcsélet újításhoz, az újdonsághoz és a fejlődéshez való sokszor ellentmondásos viszonyát.

3.1 Az újdonsághoz való viszony az ókori görög filozófiában

Az i. e. 7. századtól kezdve az i. e. 4. századig, a görögség klasszikus korszakát az emberi történelemben kivételes fejlődés, és tudományos, hadászati, politikai, valamint művészeti újítások sora jellemezte, az eredetiség és a találékonyság addig ismeretlen, kivételesen magas szintjén. A mai civilizációnk arculatát is meghatározó matematikai, orvoslási, építészeti eredmények mellett, a legjelentősebb előrelépések a kultúra területén történtek. A végbement folyamatok leírására használt *görög csoda* kifejezés a görög kultúra egyedülállóságát igyekszik kiemelni, nem véve figyelembe más ókori népek görög kultúrához való hozzájárulását. Azonban tény, hogy a klasszikus időkben léteztek olyan agyagi és szellemi feltételek, amelyek a haladás bizonyos formáinak megjelenését támogatták és ösztönözték – ezeket Ludwig Edelstein a *The Idea of Progress in Classical Antiquity* című művében részletesen elemzi, kiemelve – több korábbi szerzővel ellentétben – a görögség hitét a haladásban, bizonyítékokat szolgáltatva a progresszióra és az erre való törekvésre számos területen

(Edelstein, 1967). A görögség ciklikus időfelfogása nem tette lehetővé a mai értelemben vett lineáris és teleologikus fejlődésértelmezést, habár a technikai képességek és az ismeretek bővülésének tapasztalata nyilvánvaló volt az olyan gondolkodók számára, mint Platón, aki azonban ezt például morális és szellemi fejlődés folyamatához kötötte, és némi fenntartással viseltetett iránta, úgy véelve, hogy az újítások eltávolíthatják az embereket az igazságtól és az erkölcsi értékektől.

Godin az *Innovation and Conceptual Innovation in Ancient Greece* (Godin, 2012) tanulmányában, bemutatja innováció fogalmának intellektuális történetét, nyomon követve eredetét és fejlődését az ókori Görögországban. A szerző alapos elemzést végez arról, hogy a kulcsfontosságú görög filozófusok és írók hogyan használták és értelmezték az innováció kifejezést, négy jelentős szerzőre – Xenophónra, Platónra, Arisztotelészre és Polübioszra – összpontosítva. Jelen fejezetben számomra filozófiai hatástörténetük okán Platón és Arisztotelész munkássága kapott figyelmet, azonban Xenophón kapcsán meg kell említenem, hogy jelentős hatást gyakorol a megközelítése az innováció fogalom használatára, mivel annak elsődlegesen politikai kontextusú használata tőle eredeztethető (Godin, 2012, pp. 5-11).

Az ókori görög nyelvben az innováció fogalmára gyakran a *kainotomia* (καινοτομία) kifejezést használták, ami szó szerint újítást vagy új dolgok bevezetését jelenti. Ez a tudomány területén általánosan elterjedt volt és pozitívabb megítéléssel is bírt, mint a politikai szférában. Az újítás politikai kontextusban történő megjelenésével, legyen szó az új törvényekről, reformokról vagy államrendről, a görögség nagyobb távolságot tartott és sokszor fenntartásokkal viseltetett. Az ókori filozófusok és politikai írók szóhasználatában az innováció a *fennálló rend megváltoztatását*, egyfajta *forradalmi* tevékenységet jelentett, ami hozzájárult a fogalom vitatott jellegéhez (Godin, 2017, p. 25).

Christian Meier amellet érvel, hogy az ötödik századi görögök nem voltak a fejlődés elkötelezettei – miként nem is álltak azzal szemben – pusztán az emberi tudás és a technikai képességek gyarapodásával számoltak, elfogadták a növekedés és a sokféleség tapasztalatát. Meier a görög politika fejlődéséről írt nagyhatású művében (Meier, 1990) felvázolja a demokrácia fejlődésének útját, a reformok sorát és a társadalmi haladás eszméjének görög megfelelőjét, bemutatva, hogy az új politikai kultúra miként alakítja át a görög társadalmat, előmozdítva a szellemi fejlődést, a kritikai gondolkodást és közvetlen módon elősegítve innovációk sorát – ezért érdemes innen, a politika felől közelítenünk Platón újításhoz való viszonyához is.

3.1.1 Platón viszonya az újításokhoz

Az előzőekben leírtakhoz hasonlóan, Platón viszonya az újításokhoz kétarcú: egyszerre kritikus az innovációkkal kapcsolatban és egyben úttörő filozófiai módszerek és elméletek megalkotója. Azon műveiben, amelyekben megjelenik az újítás témaköre, a filozófus erősen szkeptikus hangot üt meg, különösképpen a politikában és a társadalmi normákban megjelenő újításokkal szemben, kiállva a meglévő struktúrák és hagyományok megőrzése mellett.

A következő szövegrészlet Platón (i.e. 427-347) Az állam című művéből származik, és a zenei és testi nevelés területén bevezetett újítások veszélyeiről szól:

„[...] az állam gondozóinak arra kell vigyázniok, hogy óvatlanul káros elferdülések ne támadjanak, és mindenekfelett azon örködjenek, hogy a zenei és testi nevelés terén ne történjenek a megállapított rendet felborító újítások, hanem a hagyományt lehetőleg híven megőrizték (...) Óvakodnunk kell ugyanis a zenei formák újító megváltoztatásától; ezzel mindent kockára teszünk, mert ha valaki a zenei formákat módosítja, ez kihat az állam legfőbb törvényeire – ezt tanítja Damón, s ezt vallom én is” (Platón, Az állam, 424 b–c, s.p.).

Platón szerint a stabilitás és a hagyomány kulcsfontosságú az oktatásban és a társadalomban, és a zenei és testi nevelés tárgykörén keresztül mutat rá, hogy a legkisebb változtatások, újítások veszélyeztethetik az állam rendjét és törvényeit, amely nézet tükrözi azon meggyőződését, hogy a társadalmi rend és harmónia szorosan összefügg a nevelés és a kultúra stabilitásával (Taylor, 1997, pp. 371–418).

A másik mű, ahol az újítás egyértelmű negatív megítélésével és kritikájával találkozunk a Törvények, amelyben Platón bemutatja a játékok példáján keresztül, hogy a permanens változások az ifjúság körében negatív következményekkel járhatnak a jellemükre és értékrendjükre nézve, ami alááshatja a társadalom alapjait. A platóni érveket ebben a szövegrészben a dialógust irányító „Athéni” vezeti elő, aki beszélgetőtársainak – Kleiniásznak és Megillosznak – a játékok és az állami törvények közötti szoros összefüggésre hívja fel a figyelmet, hangsúlyozva, hogy a játékok állandósága kulcsfontosságú a törvények stabilitása szempontjából:

„Mert, ha a játékok kérdése úgy van rendezve, hogy az örökkévalóban részesül: vagyis ugyanazon állam fiatalsága ugyanazokat a játékokat mindig ugyanúgy és ugyanazon szabályok szerint játssza, és ugyanazon játékszerekben találja kedvét, ez azt eredményezi, hogy a komoly dolgokra vonatkozó törvények is változatlanul maradnak. Ha ellenben a játékokat megbolygatják és újításokat hoznak be, és mindig újabb változtatásokkal próbálkoznak, és sohasem ugyanazt kedveli az ifjúság, akkor nem is uralkodhat állandó közmegegyezés abban a kérdésben: mi az illő és mi nem – akár a testtartásról, akár a ruházatról, akár egyéb felszerelésről

van szó –, hanem az az ifjú részesül közmegebecsülésben, aki mindig valami újat talál ki, és olyasmit hoz be, ami alakban, színben vagy más effélében eltér a megszokottól. Márpedig ennél nagyobb ártalom nincs az államra nézve.” (Platón, Törvények 7. könyv VII., s.p.).

Platón ebben a szövegben is azt állítja, hogy ha a látszólag jelentéktelen dologban is, mint a játékok, újítások történnek, az a társadalmi normák és értékek megingásához vezethet, ami végső soron az állam stabilitását veszélyezteti. A filozófus párhuzamot von a test és lélek egyensúlyának megzavarása, valamint a társadalmi rend felborulása között, kiemelve, hogy az újdonság iránti szeretet a fiatalok körében a régi dolgok leértékeléséhez és az új dolgok túlzott tiszteletéhez vezethet. Platón szerint a változás általában mind a természetben, mind az emberi viselkedésben és életmódban veszélyes, kivéve, ha rossz dolgokat változtatunk meg, illetve hangsúlyozza, hogy az emberek nehezen alkalmazkodnak az új szokásokhoz és törvényekhez, ezért a törvényhozónak arra kell törekednie, hogy fenntartsa a stabilitást, különösen az erkölcsi szabályozók terén, mert ez okozhatja a legnagyobb károkat egy állam életében. Emiatt a törvényhozók felelőssége, hogy megtalálják a módját a stabilitás megőrzésének és a folyamatos változások negatív következményeinek megelőzésének, különös tekintettel az erkölcsi értékek változásaira, amelyek a legsúlyosabb következményekkel járhatnak a társadalomra nézve (Taylor, 1997, pp. 664–685).

Platón megközelítése rávilágít az újítás kettős megítélésére: egyrészt potenciális veszélyforrás a társadalmi stabilitásra nézve, másrészt nélkülözhetetlen eszköz a tudás és a bölcsesség gyarapításában. Ez a szemlélet nemcsak filozófiatörténeti jelentőségű, de ma is releváns kérdéseket vet fel a hagyomány és a haladás egyensúlyáról, az innováció etikai vonatkozásairól, valamint a filozófia szerepéről a társadalmi változások alakításában.

3.1.2 Arisztotelész Hippodamosz államelmélet kritikája

Arisztotelész (ie.384-322) hozzájárulása az innovációról szóló filozófiai diskurzushoz sokrétű, és olyan meglátásokat kínál, amelyek a technológiai fejlődésről és az etikai megfontolásokról szóló kortárs vitákban is relevánsak. Platónhoz hasonlóan, Arisztotelész, mint innovátor is jelentős: munkássága a tudományos megismerés módozatainak kidolgozása vagy a kategóriatan területén, forradalmasította a tudás megszerzésének és szervezésének megközelítését, és szisztematikus módszerei megalapozták a modern tudományos gondolkodást.

Az Arisztotelész által bevezetett techné (τέχνη) fogalma közvetlenül kapcsolódik az innovációhoz, mivel a techné a tudásnak azt a formáját jelöli, amely konkrét készségeken és gyakorlati alkalmazáson alapul. Azonban ahogy Heidegger rámutat, a techné „nemcsak a

kézműves ténykedés és képesség neve”, hanem a „*techné a létrehozáshoz, a poiésis-hez tartozik*” (Heidegger, 1994, p. 116). A techné magába foglalja tehát az új tárgyak létrehozásának képességét, az alkotás kreatív folyamatát, azaz lehetőséget, hogy az ember új megoldásokat hozzon létre konkrét problémákra.

A Nikomakhoszi etikában az erkölcsös éltmódra nevelés kapcsán fejti ki a fejlődés lehetőségét, sőt szükséges voltát:

„Éppen ezért egyesek úgy vélekednek, hogy a törvényhozónak hívogatnia kell ugyan a polgárokat az erényre, s buzdítania kell őket az erkölcs követésére, mert hiszen azok, akiknek a fejlődése a szokás folytán már eleve erkölcsös irányba s azokkal szemben, akikhez a természet mostohább volt, büntetést és megtorlást kell alkalmaznia, sőt a javíthatatlanokat egyáltalán ki kell rekesztenie (...) Ha mármost – miként mondottuk – mindenkinek, aki erkölcsös akar lenni, helyes elvek szerint kell nevelkednie, jó szokásokat kell elsajátítania, erkölcsös életmód szerint kell élnie, és sem akaratlanul, sem akarva nem szabad a rosszat cselekednie – s ha ez csak úgy valósulhat meg, ha az ember bizonyos ésszerű elveknek és a helyes rendnek megfelelően él, amely a kellő erélyt is biztosítja” (Arisztotelész, 1987, pp. 381–382, 1180a).

A filozófus hangsúlyozza a helyes elvek, jó szokások és erkölcsös életmód fontosságát, amelyek elengedhetetlenek a társadalom fejlődéséhez. Az erényes élet nem véletlenszerűen alakul ki, hanem tudatos nevelés és iránymutatás eredménye. A szövegben Arisztotelész hangsúlyozza a nevelés, a helyes elvek és a jó szokások elsajátításának fontosságát és Platónhoz hasonlóan a törvényhozók felelősségét emeli ki a polgárok erkölcsi nevelésében. A szöveg arra utal, hogy az ember nem születik kész erényekkel vagy képességekkel, hanem ezeket tanulás és gyakorlás útján sajátítja el, az erkölcsös viselkedés nem velünk született tulajdonság, hanem olyan készség, amit fejleszteni kell – ez egyben magában hordozza pozitív újítások lehetőségét.

Arisztotelész a dolgoknak célt tulajdonító szemlélete alapján, a végső célhoz és a „jóhoz” való hozzájárulás szempontjából értékelhető minden újítás. A Nikomakhoszi etika bevezetésében megfogalmazott gondolat szerint minden emberi tevékenység valamilyen jóra irányul: *„Minden mesterség és minden vizsgálódás, de éppúgy minden cselekvés és elhatározás is, nyilván valami jóra irányul; tehát helyes az a megállapítás, hogy „jó az, amire minden irányul” (Arisztotelész, 1987, p. 5, 1094a).*

Ebből kiindulva az innovációt is annak fényében kell megítélni, hogy mennyire szolgálja az emberi jólétet és a társadalom fejlődését, azaz egy újítás akkor tekinthető értékesnek, ha összhangban van az etikai alapelvekkel, és végső soron a közösség javát szolgálja. Az innováció folyamatának tehát nem csak technikai vagy gazdasági szempontokat kell figyelembe vennie, hanem morális megfontolásokat is. Arisztotelész teleologikus

szemlélete alapján az újításokat és emberi törekvéseket nem csupán technikai vagy gazdasági érdekük alapján kell megítélni, hanem végső céljuk, erkölcsi értékük és a közjóhoz való hozzájárulásuk szempontjából is, ami egy átfogóbb etikai dimenziót ad értékelésükhöz.

A történelmi korszak, amelyben Arisztotelész alkotott, számos szempontból nyitottabb volt az újításokra, ami a filozófus munkáiban is tükröződik. Arisztotelész megsem tekinthető azonban elkötelezett fejlődéspártinak, több helyen erős kritikát is megfogalmaz újítókról, újításokról. Erre kiváló példa a Politika második könyvében az államelméletek kapcsán kifejtett kritikája Hippodamoszról, aki az i. e. ötödik században élt milétoszi származású várostervező volt, és akinek kormányzási és városszervezési elméleteit a gyakorlatban is alkalmazták Pireuszban és Rhodoszban.

Lewis Mumford amerikai történész-filozófus, aki gyakran foglalkozott az emberi értékek és a technológiai fejlődés közötti konfliktusokkal, valamint az urbanizáció társadalmi következményeivel, a talán legismertebb *The City in History* című művében tárgyalja Hippodamosz várostervezéshez való hozzájárulását. Hippodamosz felismerte, hogy a város formája tükrözi a társadalmi rendet és úgy vélte, hogy egy város átalakításához megfelelő társadalmi szerkezeti változásokra van szükség. A várostervezést gyakorlati feladatnak tekintette, amely egy racionálisabb társadalmi rendet képes kialakítani. Úttörő szerepet játszott az ókori görög városrendezésben, innovatív módon összekapcsolva a városok fizikai elrendezését és társadalmi struktúráját (Mumford, 1961, pp. 172–173).

Hippodamosz Arisztotelész általi megítélése kétségtelenül rossz: azt sugallja, hogy a *hübrisz* uralkodott el rajta, olyan területeken nyilvánít magabiztosan véleményt, amihez nem ért, és magát az átlagos polgár felé helyezi. Arisztotelész összefoglalja Hippodamosz elméletét az ideális városállamról, amelyet tízezer kormányzó polgár alkot, akik kézművesekre, földművesekre és katonákra oszlanak, ők közösen választják vezetőiket. A földtulajdon is három részre oszlik ebben az államban: köz-, magán- és szent földre. Hippodamosz társadalmi újításokat is bevezetett „*törvényt hozott, hogy megbecsülés járjon azoknak, akik a városállamra hasznos találmányok szerzői, és arra is, hogy a háborúban elesettek árvái közpénzen tartassanak*” (Arisztotelész, 1969, p. 22, 2/8.)

Arisztotelész az elsődleges kritikáját Hippodamosz elméletéről a polgársághoz kapcsolódóan fogalmazza meg: habár mindenkinek juttat a város területéből az egymáshoz való viszonyaikat és a társadalomban betöltött szerepüket nem veszi figyelembe, ami a közösség vezetése kapcsán feszültségeket és elégedetlenséget szül. A második kritika a törvénykezési részt illeti. Itt fő problémaként azt azonosítja, hogy az esküdtek bíraskodási joga a bírói hivatal

szerepét és jelentőségét teszi zárójelbe. Szükségesnek látja valamilyen felsőbb tekintély bevonását (mint amilyen például egy vének tanácsa) amelyhez a vitás kérdésekben fellebbezni lehet. és ami Hippodamosz államában hiányzik. Számunkra a legérdekesebb azonban a harmadik kritikája, amely az újításokra vonatkozik. *„Vajon káros vagy hasznos-e a városállamnak, ha ősi törvényeit megbolygatják, mikor valami jobb kínálkozik?”* – teszi fel a kérdést. Amikor Hippodamosz azon javaslatát említi, ami a városnak hasznos újítások bevezetőinek és feltalálóinak juttatandó jutalomról szól, Arisztotelész a hagyományra és a törvényre utal és Hippodamosz tájékozatlanságát ostromozza, aki nem rendelkezik kellő tudással és abban a hiszemben próbál valami újat kezdeményezni: *„hogyan ez másutt még nincs törvénybe iktatva; pedig hát ez a törvény Athénben és más államokban is megvan”* (Arisztotelész, 1969, p. 22, 2/8.)

Arisztotelész is érzi ezen a ponton, hogy határozott szembehelyezkedése általánosságban magával az újítással félreérthető lehet, ezért több irányból igyekszik a Hippodamosz szembeni kritikának ezt az életet tompítani. A törvények kapcsán ekként ír: *„még az írott törvényeket sem tanácsos változatlanul hagyni (...) egyes törvényeket bizonyos esetekben meg kell változtatni”* (Arisztotelész, 1969, p. 24, 2/8.). Ezek a gondolatok egyértelműen egy fejlődés mellett elkötelezett gondolkodó képét mutatják, azonban a szakasz végén kiemeli, hogy mindezek ellenére *„más szempontból a változtatás, úgy látszik, nagy óvatosságot igényel”* (Arisztotelész, 1969, pp. 24-25, 2/8.).

A Hippodamoszt célzó némileg gúnyos kritikája miszerint *„ő volt az első, aki anélkül, hogy politikus lett volna, a legjobb alkotmányról véleményt próbált nyilvánítani”* (Arisztotelész, 1969, p. 23, 2/8.). – egy idealista szakember kritikája, akin Arisztotelész a polisz iránti elköteleződést és előzetes politikai tapasztalatot kéri számon, ezek hiányában vitatva jogát a város ügyeibe való szakértő bevonódásba. Küllemének és viselkedésének sajátosságait az elméleteire is rávetítve, azok különbségét és valóságtól elrugaszkodott voltát igyekszik hangsúlyozni, ezzel érvényességüket és működőképességüket is kétségbe vonva.

Hippodamosz társadalmi reformjai, amelyek az egyes osztályok közötti különbségek elmosásával az osztályok közötti feszültség csökkentését célozzák, a földek státuszának rendezése, a szociális innovációk, mint pl. az árvaellátás átszervezése, az innovációs tevékenységek előmozdító díjak kiírása, a város szerkezetének újragondolása olyan kezdeményezések, amelyek az akkori politikai és társadalmi keretek radikális újrarajzolását jelentik. Arisztotelész távolságtartása nem az újításra, mint általános tevékenységre vonatkozik (ezt ki is emeli például az orvoslás vagy a testedzés kapcsán), hanem a törekenynek gondolt

társadalmi egyensúly és a hagyományok (amelyek a törvényeket és viselkedési szabályokat is magukban foglalják) garantálta kohézió megbomlásának fenyegetése lebeg előtte. Arra mutat rá, hogy tévedés azt feltételezni, hogy csak azért, mert az újítást az egyik környezetben jó dolognak tekintik, az egy másik környezetben is előnyös lesz és a politika területén ez nem így is működik, ahogy fogalmaz: „*a meglevő törvényeknek más, új törvényekre való könnyelmű átformálása tulajdonképpen a törvény erejének gyengítése*” (Arisztotelész, 1969, p. 25, 2/8.).

3.2 Az újítás szerepe a középkori arab filozófiában

Mielőtt a nyugati filozófiatörténet időfolyamában tovább haladunk, érdemes egy kitérőt tennünk az iszlám filozófia felé, ahol az újításról, az innovációról (arabul: *bi'dah*) való gondolkodás első (vallás)filozófiai ihletettséggű, de már szisztematikus elméletét megtaláljuk. Az iszlám tudományosság, különösen a középkori arab filozófia sok esetben alulértékelt szerepet játszik, a nyugati filozófia fejlődésében betöltött szerepéről hajlamos elfeledkezni a történetírás (de Boer & Jones, 1933, p. 29). Letagadhatatlan azonban, hogy maga az iszlám civilizáció fontos közvetítő szerepet töltött be az antik filozófiai műveltség és a középkori nyugati filozófiai gondolkodás között. Az olyan filozófusok, mint al-Farábi, Avicenna vagy Averroes munkássága jelentős hatást gyakorolt a skolasztikára, többek között azáltal, hogy ezek a tudósok megőrizték és továbbadták az ókori görög filozófia számos művét, illetve egyben hozzá is járultak azok értelmezéséhez és egyes elméletek továbbfejlesztéséhez, különösen a racionális gondolkodás és a vallási tanítások összeegyeztetése kapcsán. Emiatt is érdemes megvizsgálni az újítás és fejlődés gondolatát az iszlám filozófiában, mivel úgy gondolom, hogy ennek tanulmányozása nélkül nem lehet teljes képet alkotni a fogalmak evolúciójáról – habár kijelenthető, hogy a nyugati filozófiára ez az újítás-koncepció közvetlen hatást nem gyakorolt.

Az iszlám filozófia aranykorának első jelentős filozófusát, al-Farábit (kb. 872-950), elsősorban a görög filozófia és az iszlám vallási hagyomány összeegyeztetésére kísérletet tevő munkásságáról ismerjük. Igyekezett szintetizálni az eltérő filozófiai hagyományokat, hangsúlyozta az ész és az empirikus megfigyelés fontosságát a világ megismerésében (Maróth, 2007, p. 135) és a filozófiában eszközt látott a társadalom javítására, ezért a társadalom vizsgálata számára a filozófiai vizsgálódás nélkülözhetetlen részét képezte. Az eszményi város lakosainak nézeteiről című írásában fontos szerepet kap a tudás és a cselekvés az *eszményi* és az *ostoba, tévelygő* városok megkülönböztetésében (Maróth, 2013). Az eszményi város lakóinak közös jellemzőiről szóló fejezetben a tudás és megértés fokozatos fejlődését hangsúlyozza, kiemelve, hogy az emberek különböző szinteken képesek felfogni az igazságot: vannak, akik kutatják a helyes utat és magasabb szintű megértésre törekszenek, ami a

folyamatos fejlődés és az igazság keresésének gondolatát tükrözi (al-Farábi, 1997, 34. fejezet). al-Farábi számára az etikai tudás fizikai és metafizikai alapokon nyugszik, hangsúlyozza az elméleti tudás szerepét, mint az erény elengedhetetlen előfeltételét (Mattila, 2017, pp. 124–137). A filozófust úgy határozza meg, mint aki ötvözi az elméleti tudást az erénnyel és a politikai képességgel. Az eszményi városban a lakosok erényről alkotott közös véleménye végső soron filozófiai ismereteken és a gyakorlati gondolkodás fejlett képességén alapul. A fejlődés gondolata itt elsődlegesen egy erkölcsi jellemformálásként jelenik meg, amely az erkölcsi és intellektuális tökéletesedés felé haladásban ölt testet. al-Farábi optimista módon tekint az emberi potenciálra, úgy látja, hogy az egyén és a társadalom egyaránt képes a morális, intellektuális és spirituális fejlődésre, fokozatosan haladva a hétköznapi erényfelfogástól a filozófiai–vallási erény felé.

A korábban már említett – szó szerint újítást jelentő – *bi'dah* fogalom rávilágít a fennálló hagyományok, kulturális kontextusok és az újítások közötti feszültségre. Eredendően egy negatív jelentéstartalommal bír az iszlám vallásban, azonban ez nem feltétlenül jelenti a világi újításokkal szembeni ellenszegülést is: maga az iszlám civilizáció izgalmas ebből a szempontból, mivel a vallási újítások elvetése mellett a tudomány és a technológiai fejlődés vívmányainak elutasítása nem általánosan jellemző. A folyamatos egyensúly keresése az általános inkább, a társadalmi és politikai haszonnal bíró új javak, modellek és eljárások, valamint a vallási hagyományok között. Vallásjogi értelmében a *bi'dah* a hit, a vallási gyakorlat vagy szertartás bármely olyan megváltoztatását jelenti, amely a Koránban és a Hádiszban – a Mohamed prófétához kapcsolódó hagyományok gyűjteményében – szereplő elvektől eltér. A sáfiita vallásjogi iskola alapítója, al-Sáfi'í imám (meghalt 820) nézeteit követi a legtöbb jogtudós, aki különbséget tesz az újítások között, mint a *bid'ah hasanah* (jó vagy ajánlott újítás), amely a vallási gyakorlat új és hasznos kiegészítéseire utal – szigorúan a szövegek és tanítások alapján, olyan technológiai, orvosi vagy társadalmi fejlesztések, amelyek összhangban vannak az iszlám alapértékeivel és elősegítik a közösség javát. Illetve a *bid'ah sayyiah* (elítélendő vagy kifogásolható újítás) amelyeket nem támasztanak alá az iszlám szent szövegei, abból nem vezethetők le, eltérnek a tanításoktól és alapvetően emberi spekulációk és törekvések eredményei.

Ezt a megkülönböztetést elfogadta al-Gazáli (1058-1111) is, aki az iszlám egyik legbefolyásosabb középkori gondolkodója volt, és a klasszikus narratíva (Cohen, 1994; Nasr, 1968) szerint leginkább a filozófiával való szakítás egyik vezéralakjaként ismert az iszlámban.

Az egyik legismertebb munkája, ahol az újítás témáját tárgyalja, az „*Ihya' 'Ulum al-Din*” (A vallástudományok új életre keltése), ebben így fogalmaz:

„*As regard to what is claimed as innovated after the death of the Prophet (saw), in fact, not every innovated act is legally prohibited. However, what is prohibited is the innovation disagrees with the established Sunnah or abolishes a legal command despite the existence of its underlying cause. Innovation, on the contrary, might have been obligated in some cases when the underlying cases changed.*”¹⁵ (al-Gazáli, 1993, p. 3.).

al-Gazáli hangsúlyozta annak fontosságát, hogy az újításokat a Koránnal, a Mohamed tanításait útmutatásait és cselekedeteit őrző hagyománnyal – a Szunnával – és az iszlám alapelveivel való összhangjuk alapján értékeljük. Kritizálta minden újítás vak elutasítását, mivel úgy vélte, hogy ez stagnáláshoz vezethet, és ehelyett egyensúlyt igyekezett teremteni az iszlám alapvető tanításainak megőrzése, valamint a szükséges és hasznos újítások engedélyezése között. Ezt a nézőpontot tükrözta szélesebb körű filozófiai megközelítése, amely a hagyományos iszlám gondolkodást és hitet igyekezett összeegyeztetni a racionális vizsgálódással és gyakorlati megfontolásokkal, ösztönözve a kritikai gondolkodást. al-Gazáli a spirituális önéletrajzi írásában (Megszabadulás a hibáktól – al-Munqidh mina 'd-Dalal) egy újszerű megközelítést javasol a filozófiai tanok értékelésére (al-Ghazáli, 2011). Ahelyett, hogy kategorikusan elutasítaná vagy elfogadná a *falászifa*, azaz a filozófusok eszméit, egy árnyaltabb szemléletet ajánl és arra biztatja az olvasóit, hogy minden egyes filozófiai gondolatot alaposan és elfogulatlanul vizsgáljanak meg csakúgy, mint az ész és az érzéki tapasztalat helyességét. Véleménye szerint egyes filozófiai tanok valóban ellentmondhatnak az iszlám alapvető tanításainak, és ezeket természetesen el kell utasítani, ugyanakkor hangsúlyozza, hogy számos filozófiai eszme összhangban lehet az iszlám elveivel, sőt, gazdagíthatják is a muszlim gondolkodást és közösséget. Szerinte, ha egy gondolat racionális, bizonyítható, és nem mond ellent a Koránnak vagy a Szunnának, akkor azt érdemes megfontolni, függetlenül annak eredetétől (Griffel, 2009, p. 120). al-Farabi – hoz hasonlóan ő sem veti el az erkölcsi fejlődés lehetőségét, elképzelése szerint az emberi természet kezdetben hiányos és elégtelen, és az erényes jellemvonások kifejlesztéséhez komoly erőfeszítésre, folyamatos fegyelmezésre és képzésre van szükség. Ez a fajta nyitottság és kritikus gondolkodás al-Gazáli munkásságának egyik legfontosabb és legmaradandóbb öröksége.

¹⁵ „*Ami a Próféta (béke legyen vele) halála után újításként bevezetett dolgokat illeti, valójában nem minden újítás tilos jogilag. Azonban az olyan újítás, amely ellentmond a megalapozott Szunnának, vagy eltöröl egy jogi parancsot annak ellenére, hogy annak alapvető oka továbbra is fennáll, tiltottnak minősül. Ezzel szemben az újítás bizonyos esetekben akár kötelezővé is válhat, amikor az alapvető körülmények megváltoznak.*” (saját fordítás)

Abu Bakr ibn Tufajl (kb. 1105-1185) a nyugati iszlám filozófiai hagyomány utolsó nagy alakja, aki A természetes ember címet viselő allegorikus szövegében meséli el Hajj ibn Jakzán történetét, aki egy szigeten egyedül felnőve, a környezetének megfigyelésén keresztül kezdi feltárni a világ titkait, a tudomány, filozófia és erkölcs alapelveit. A történet szerint az ifjú képes új megoldások, eszközök kitalálására a változó környezeti körülményekhez való alkalmazkodva, illetve nyitott az új tapasztalatok, jelenségek megértésére, folyamatosan törekedve az egyéni tökéletesedésre. A mű azt mutatja be, hogyan jut el az ember pusztán a saját megfigyelései és gondolkodása révén az isteni igazság felismeréséhez, anélkül, hogy valaha is kapcsolatba került volna vallással vagy társadalommal. Ibn Tufajl ezzel azt kívánja demonstrálni, hogy az emberi értelem képes önmagában is eljutni a legmagasabb szintű filozófiai és spirituális felismerésekhez, képes felfedezni a természet és vallás igazságait kinyilatkoztatás nélkül is. Hajj felfedezései összhangban vannak a kinyilatkoztatott vallás tanításaival, és kiegészítik azokat, a természetről, gondolkodásról és Istenről alkotott nézeteit a hit tételei is alátámasztják (Tufajl, 1963). A könyv explicit módon hirdeti az innovatív gondolkodást, a vallási és társadalmi újítások lehetőségét és szükségességét. Központi üzenete, hogy az emberek különbözőek, és mindenkinek joga van a saját útját járni az igazság keresésében. A szöveg hangsúlyozza az erőszakmentesség és a békés együttélés fontosságát, függetlenül a vallási vagy ideológiai különbségektől, és óva int a mások nézeteinek erőszakos megváltoztatására tett kísérletektől, mivel ezek veszélyeztethetik a társadalmi harmóniát (Attar, 2007, pp. 57–58.). Tufajl európai hatástörténetének Samar Attar *The Vital Roots of European Enlightenment* címmel egy teljes könyvet áldoz, nem kevesebbet állítva mint, hogy az európai felvilágosodás gondolkodóinak munkásságában visszaköszönnek az andalúzi filozófus elméletei. Ezeket az állításokat többen kritizálják, rámutatva, hogy óvatosan kell kezelni az olyan állításokat, amelyek közvetlen kapcsolatot feltételeznek az arab és az európai irodalom között konkrét bizonyítékok nélkül, kiemelve Attar megközelítését, mert az gyakran bizonyítatlan állításokra és homályos párhuzamokra támaszkodik (Kruk, 2011).

Az iszlám filozófia újítás–konceptiójának vizsgálata al-Farábi, al-Gazáli és Ibn Tufajl munkásságán keresztül, rávilágít arra, hogy az iszlám civilizáció képes volt egyensúlyt teremteni a hagyományörzés és az intellektuális fejlődés között. Bár az európai felvilágosodásra gyakorolt hatásuk mértéke vitatott, az vitathatatlan, hogy ezek a gondolkodók olyan szellemi örökséget hagytak maguk után, amely ma is inspirálja a kultúrák közötti párbeszédet és a kritikus gondolkodást, példájuk arra tanít, hogy az igazi innováció nem a múlt elvetésében, hanem annak kreatív újraértelmezésében rejlik.

3.3 Aquinói Szent Tamás a törvények megváltoztatásáról

Az előző fejezetben tárgyalt iszlám filozófiai hagyomány, fontos alapokat teremtett a hit és az értelem összeegyeztetéséhez, amely később a keresztény skolasztika központi elemévé vált, és eklatáns módon tudjuk Aquinói Szent Tamás (1225–1274) munkásságában megragadni. Bár az iszlám filozófia Szent Tamás bölcséletére gyakorolt közvetlen hatása tudományos vita tárgyát képezi (Corbin, 1993) a vallási tanítás és a filozófiai értelmezés összehangolására irányuló megközelítéseikben figyelemre méltó párhuzamok vannak.

Aquinói Szent Tamás fő művében, a Summa Theologiae-ben kísérletet tett – a kinyilatkoztatás elsőbbségének fenntartásával – a racionális gondolkodás integrálására a teológiába, egy koherens világkép megalkotására törekedve. Vanyó László kiemeli, hogy *„a kereszténységnek a görög gondolkodással szemben a bölcsélet síkján maradvá – egy olyan gondolkodási formált kellett érvényesítenie, amelyben az értelmezés látóhatárát az ember képezi, és amelyben tartalmilag minden istenközpontú”* (Vanyó, 1970, p. 157). A fő különbség abban rejlett, hogy míg a görög gondolkodás kozmocentrikus volt, addig a keresztény nézőpont az embert helyezte középpontba. Szent Tamás filozófiájában az ember vált a létezés értelmezésének kiindulópontjává, ami lehetőséget adott a világ és Isten újfajta megközelítésére. Ez az antropológiai fordulat megalapozta a későbbi nyugati gondolkodás fejlődését, különösen az egyén szabadságának és felelősségének hangsúlyozását – többek között emiatt is érdekes Szent Tamás filozófiája a tárgyunk szempontjából.

A már említett Summa Theologiae-ben, a törvények változásáról és módosításáról szóló XCXVII. kérdésében találkozunk az újítás kapcsán kifejtett legteljesebb gondolataival a filozófusnak. Ebben a szövegrészben a fő állítás az, hogy bár a törvényeknek bizonyos mértékig állandónak kell lenniük, a változó körülmények miatt néha szükség lehet a módosításukra:

„ha nem változtak volna az emberi törvények, amikor jobbak kínálkoztak, sok nem kívánatos dolog következett volna be, mivel úgy találjuk, hogy az antik törvények sok kezdetlegességet tartalmaznak. Tehát úgy látszik, hogy a törvényeket meg kell változtatni mindannyiszor, amikor helyettük valami jobbat lehet alkotni” (Aquinói Szent Tamás, 2011, p. 54).

Azonban a törvények megváltoztatása nem történhet könnyedén vagy gyakran –állítja a szerző –, mivel ez alááshatja a törvények tekintélyét és betartását. A törvényeket csak akkor szabad megváltoztatni, ha a változtatás egyértelműen a közjót szolgálja, és a várható előnyök felülmúlják a változtatással járó hátrányokat. A törvények stabilitása és a megszokás fontos szerepet játszik azok betartásában, ezért óatosan kell bánni a módosításokkal. A szöveg különbséget tesz a törvények és más emberi alkotások (pl. mesterségek) között, ezekkel sokkal

megengedőbb: „amikor az ész számára jobb megoldás kínálkozik, meg kell változtatni azt, amihez korábban ragaszkodtak” (Aquinói Szent Tamás, 2011, p. 55). Míg az utóbbiak esetében könnyebben lehet újítani, a törvények esetében nagyobb körültekintésre van szükség. Szent Tamás szerint a törvények megváltoztatása csak akkor indokolt, ha jelentős haszonnal jár vagy rendkívül szükséges például, ha a régi törvény erkölcsileg kifogásolható vagy káros hatású: „törvényeket meg kell változtatni, de nem akármilyen jobbítás érdekében, hanem nagy hasznosság végett, vagy a szükségesség miatt” (Aquinói Szent Tamás, 2011, p. 55).

A szöveg a törvények stabilitása és a szükséges változtatások közötti egyensúly fontosságát hangsúlyozza, szem előtt tartva a közjó szolgálatát. Alapvetően a szövegből inkább az újítások elfogadás melletti állásfoglalás olvasható ki, elismeri azok szükségességét, de óvatosságra is int velük kapcsolatban. Ez a gondolat jól illeszkedik Szent Tamás komplex emberképéhez, amely egyensúlyt teremt az ember racionális, társadalmi és történelmi dimenziói között.

3.4 Az innováció, mint az adaptáció eszköze Machiavelli államelméletében

Niccolò Machiavelli-t (1469-1527) elsősorban, mint a modern politikai gondolkodás egyik előfutárát ismerjük, aki a történelmet és a politikát egy pragmatikus és realista megközelítésben vizsgálta. Benoit Godin rámutat, hogy annak ellenére, hogy az innováció kifejezés ritkán fordul elő a műveiben (Godin, 2015a, pp. 58-75), munkássága mégis kulcsfontosságú az innováció fogalmának történeti fejlődésével kapcsolatban. Machiavelli számára az innováció az uralkodás, a hatalom megszilárdításának és a változó körülményekhez való alkalmazkodásnak az eszköze. Esetében is a társadalmi rend és a törvények megújulása és megváltozása/megváltoztatása az innováció értelmezésének a kulcsa, csakúgy, mint a korábbi korszakok filozófiai gondolkodásában. Munkáiban az államszervezet megújítását célozza, az újítást alapvetően pozitív és szükségszerű folyamatnak értékelve, de hangsúlyozva a megelőző korszakok hagyományainak jelentőségét és az összhang keresését.

Legjelentősebb művében, a filozófus halála után 1532-ben megjelent *A fejedelem*-ben az innováció, mint új törvények és intézmények bevezetése jelenik meg, amely az uralkodót a hatalom megszerzésében és megtartásában segíti. Machiavelli kiemeli: a fejedelemnek rugalmasnak kell lennie, hogy mindig képes legyen alkalmazkodni a változó körülményekhez – az innováció az adaptáció egyik fontos eszköze, amely sok esetben erőszakot is megkövetel. A VI. szakaszban jegyzi meg, hogy az új törvények és rendelkezések bevezetése az egyik legnehezebb feladat egy új uralkodó számára, mivel ellenállásba ütközhet mind a régi rendszer haszonélvezői, mind az új rend bizonytalan támogatói részéről. A szerző szerint a sikeres

változtatásokhoz gyakran erő alkalmazása szükséges, és „*a fegyveres próféták mindannyian győztek, a fegyvertelenek pedig elbuktak*” (Machiavelli, 1979, p.23). Ebben a kontextusban az erőszak nem feltétlenül nyílt fizikai agresszióra utal, inkább egy határozott hatalomgyakorlási modellt sugall, amely az újítás mellett elkötelezett támogatókat is maga mögött tud. Az „(...) *emberek inkább a jelenben élnek, mint a múltban; és ha jól érzik magukat a jelenben, azt élvezik, s mást nem keresnek; és a fejedelmet védelmükbe veszik, ha más dolgokban megállja a helyét*” – hívja fel a figyelmet később arra (Machiavelli, 1979, p. 79.), hogy a bölcs és erényes cselekvés révén, az új uralkodó biztosabb helyzetbe kerülhet a korábbiaknál, mert nagyobb figyelem irányul a tetteire és képes jobban megragadni az alattvalók figyelmét, elnyerve a támogatásukat. Az uralkodónak, mint politikai újítónak, szüksége van az erőszak mellett ravaszságra, a színlelés képességére is, hogy fenntartsa hatalmát a még nem stabilizálódott környezetben, és állandóan szembe kell néznie a változó körülményekkel, illetve az emberek kiszámíthatatlan viselkedésével, emiatt a gyors cselekvésre való képesség is létfontosságú (Pocock, 1975, pp. 156-181).

A modern történelemfilozófia egyik megalapozásának tekintett *Beszélgetések Titus Livius első tíz könyvéről* című művében, Machiavelli kevésbé radikális nézőpontot fogalmaz meg az újítással kapcsolatban, inkább a megújulás, a régi alapokhoz való visszatérés jelentőségét hangsúlyozza. Már a Fejedelemben is többször említi, és potenciális veszélyforrásként hivatkozik rá, hogy az emberek sokszor idegenkednek az új dolgoktól – „*nem nagyon bíznak az új dolgokban biztos tapasztalat nélkül*” (Machiavelli, 1979, p. 23). Emellett az újítás sokak számára mégis vonzó lehetőségként mutatkozik meg:

„az emberek áhítoznak az újdonságra, mégpedig inkább azok kívánnak változást, akik jó sorban élnek, és kevésbé azok, akik rossz viszonyok között élnek, mert – mint erről más helyen már beszéltem – az emberek beleunnak a jóba, a balsors viszont megtöri őket. Ennek a vágyódásnak következtében tárt ajtó fogad mindenkit, aki változást akar hozni az országba: ha idegen, követői támadnak, ha országbéli, köréje tömörülnek, növelik a hatalmát, előnyökben részesítik, tehát akárhogy is viselkedik, mindenképpen siker koronázza tevékenységét.” (Machiavelli, 1979, p. 387-388).

A Beszélgetésekben óvatosságra int, például az első könyv IX. szakaszában – ahol az államalapítás, illetve az állam újjászervezésének kérdéskörét tárgyalja. Itt megfogalmazza tézisét, amely szerint az állam hatékony megszervezéséhez vagy megreformálásához elegendő (és szükséges) egyetlen erős vezető, aki teljhatalommal rendelkezik – azonban a hatalom megtartása és az állam hosszú távú stabilitása már sokak közreműködését igényli. Fontos ezért, hogy aki az alattvalók széles körének elégedettségét kívánja kivívni, az nem eszközölhet

radikális változásokat, az ősi szokásokat meg kell őriznie és be kell építenie az új hatalmi struktúrákba: „*a megújulás a kezdetekhez való visszatérést jelenti*” (Machiavelli, 1979, p. 326). Az idők folyamán az intézmények és értékek romlásnak indulhatnak, ezért szükség van a rendszeres megújulásra, hogy fenntartsák az eredeti erényeket és rendet. Ez a megújulás történhet külső események hatására vagy belső szándékból eredeztetve, például új törvények alkotása vagy egyéni példamutatás révén. Machiavelli Rómát említi példaként, ahol külső támadások és belső reformok által sikerült megújulni és visszatérni az eredeti alapelvekhez, ami hosszú távú stabilitást és fejlődést biztosított az államnak.

Ahogy Godin megjegyzi, Machiavelli számára az innováció egy stabilizáló erő, nem valamilyen forradalmi változások eszköze (Godin, 2014a, p. 21) és a folyamatosan változó világban az egyensúly megőrzése miatt szükséges. Szemben az antikvitás gondolkodóival, Machievellinél az innováció – a társadalmi rend megváltoztatása – nem kap semmilyen negatív jelentéstöbbletet, lehetőség és szükségszerűség az állam jobbá tételére, és a hatalom megőrzésére. Christopher Holman tovább megy, szerinte az epikureus filozófiai hagyomány hatására Machiavelli egy olyan világképet vázol fel, amely alapjellemezője a kontingencia, minden kaotikus és bizonytalan, semmi sem rendelkezik állandó, megváltoztathatatlan lényeggel vagy természettel (Holman, 2018, pp. 43-75). Ebben a transzcendens alapot elutasító szemléletben, egy olyan világban helyezi el az embert, ahol a társadalmak és politikai rendszerek folyamatos változásban vannak és az egyének a vágyaiktól és törekvéseiktől vezérelten szabadon alakíthatják önmagukat és társadalmi környezetüket. Ez a felfogás az embert kreatív, az újításra nyitott lényként mutatja be, aki képes alakítani saját sorsát és környezetét. Machiavelli felfogása szerint az emberi vágyak és érdekek sokfélesége elkerülhetetlenül konfliktusokhoz vezet, ami paradox módon nem negatívum, hanem a társadalmi és politikai innováció motorja lehet, ösztönözve az új megoldások és struktúrák kialakítását a folyamatosan változó körülmények között. Machiavelli ezen olvasata a világ és az emberi természet folyamatos változásáról és az új valóságok megalkotásának lehetőségéről, szorosan kapcsolódik a modern innovációfogalomhoz, amennyiben mindkettő hangsúlyozza a kreativitás, az alkalmazkodóképesség és az újítás központi szerepét a fejlődésben, legyen szó a politikáról, társadalomról vagy a tudományról.

3.5 Az új dolgok hasznosságáról: Francis Bacon az innovációról

A skolasztikus gondolkodás válsága a XIV. századra nyilvánvalóvá vált, az értelem és a hit viszonyrendszerében egyre inkább a szétválasztás, a távolodás érvényesült. Az újítás kérdésköre a racionális kutatások és a tapasztalati alapú tudományos megközelítés előtérbe

kerülésével növekvő jelentőséget kapott, elsőként a természetfilozófiát, a megismerést és a tudományos módszereket új alapokra helyezni szándékozó Francis Bacon (1561-1626) munkásságában.

Az 1605-ben megjelent *The Advancement of Learning* címet viselő művében a tudás és a tanulás természetéről értekezik, itt említi meg azokat a problémákat, amelyek a tanulásban nehézségként jelentkezhetnek a hagyományokhoz való ragaszkodás és az újdonságok elfogadása kapcsán. Úgy véli, az emberek hajlamosak két szélsőség felé hajlani: vagy a régi (esetében az antikvitásból eredeztetett) tudást részesítik előnyben, vagy pedig a teljesen újat és modernet. Rámutat, hogy az idő múlásával a régi és az új ismeretek gyakran konfliktusba kerülnek egymással, a régihez való ragaszkodás gyakran akadályozza az új fejlesztéseket, míg az újdonság hajlamos lerombolni a régit ahelyett, hogy arra építene. A szerző szerint érdemes időt szánni a múlt tanulmányozására, de nem szabad ott megrekedni, a cél az, hogy a múlt tanulságait felhasználva haladjunk előre, ezért arra buzdít, hogy találjunk egyensúlyt a múlt tisztelete és az újítás között, és egy kiegyensúlyozottabb, reflektívebb megközelítést javasol a tudás fejlesztésére (Bacon, 1901, p. 62.).

Bacon korai művei közé tartozik az 1609-ben megjelent, *De sapientia veterum* (Az ősök bölcsességéről) című mű, amely görög mitológiai történetek filozofikus–allegorikus értelmezéseit tartalmazza. A könyvben szereplő Daidalosz, avagy a mesterember címet viselő írásában a görög mitológiai alakot használja fel az emberi kreativitás és mesterségbeli tudás kettős természetének bemutatására. Daidalosz beemelése nem véletlen, egyszerre példázza az emberi alkotóképesség csodálatra méltó, és határokat nem ismerő, veszélyes oldalát. A szöveg rámutat a mesterségbeli tudás ambivalens természetére: bár Daidalosz alkotásai hozzájárultak az istenek dicsőítéséhez és a városok szépítéséhez, ugyanakkor veszélyes és erkölcstelen célokat is szolgáltak, mint a példaként felhozott mitikus történetben a szerkezet, amelyet a kréti királyné Pasziphaé kérésére készít a mester, aki így kielégíthette egy bika iránt támadt vágyát, létrehozva a Minotauruszt. A Daidalosz alkotta labirintus *„melynek rendeltetése és célja káros, terveit és kivitelezése azonban kiváló és csodálatra méltó”* (Bacon, 2019, p. 224), egyszerre ördögi eszköz az áldozatok csapdába ejtésére, másfelől a funkciója egyben az is, hogy az emberek világtól elzárva tartsa a szörnyeteget, ezzel ellensúlyozva Daidalosz rossz tetteinek következményét. Ez a kettősség világít rá arra, hogy az emberi találékonyság és kreativitás, az erkölcsi megítélés szempontjából nagyon eltérő eredményeket hozhat, illetve egy még fontosabb kérdést is felvet: az alkotó személyes felelősségét, különös tekintettel a saját tevékenysége következményeinek kezelésében (Bacon, 2019, pp. 223–225).

Bacon ebben az allegorikus szövegben azt mutatja be, hogy a tudományos ismeretek és a technológiai innováció révén megszerzett hatalom jelentős politikai következményekkel és egyúttal erkölcsi dilemmákkal is jár. Aggodalmát fejezi ki azzal kapcsolatban, képesek vagyunk-e ilyen hatalmas erőket pusztán jó célokra irányítani, és megkérdőjelezi, hogy a törvények, az oktatás vagy más intézkedések hatékonyan tudják-e biztosítani, hogy a tudományos fejlődést etikus és az emberiség javára használják fel (Studer, 1998, pp. 219–222.).

A tudomány társadalmi hasznának optimista ábrázolása ellenére Bacon tisztában volt a tudományos hatalommal való visszaélés lehetőségével. A technika fejlődése jelenthet veszélyt az emberiségre, ezt nem képes ellensúlyozni azon tudósok eredendően jóra irányuló szándéka sem, akik ezt a folyamatot ellenőrzik, mert egyes tudósok, akiket a tudás és esetleg a személyes hírnév keresése vezérel, nem mindig a közjó érdekében cselekszenek. Bacon egy olyan új, a hatalmi viszonyokat újradefiniáló tudomány mellett szállt síkra, amely egyértelműen az egész emberiség javát szolgálja. Bacon megközelítése az volt, hogy olyan tudományos közösséget támogasson, amelyet világi és erkölcsi elvek egyaránt vezérelnek, hogy a tudományos fejlődésben rejlő teljes potenciált a társadalom javára fordítsa. Úgy vélte, a tudományos és technológiai innováció párbeszédet tud kezdeményezni és együtt tud létezni a hagyományokat tisztelő erkölcsi, vallási és politikai erőkkel, illetve ésszerű és mérsékelt reformokhoz vezethet. Itt világlik ki legélesebben Bacon (jegyezzük meg: és a korábbi korszakok, gondoljunk csak Platón vagy Szent Tamás filozófiájára) innováció értelmezésének sajátossága, amiben az újításnak a politikai jellege kap elsődleges fókuszot és fő tárgyalása is a hatalom kérdéskörének kontextusban történik (Paterson, 1989, pp. 457–463).

„Alapítványunk célja az, hogy megismerjük a természetben az okokat, mozgásokat és belső erőket, s az emberi hatalom határait előbbre vigyük, ameddig csak lehet” – foglalja össze a Salamon Házának atyja a tevékenységüket az Új Atlantisz című írásában (Bacon, 2001, p. 17). A Salamon Háza egyfajta tudományos kutatóintézetként funkcionál Bansalem utópisztikus társadalmában, amelynek feladata, hogy az emberek javát szolgáló felfedezések révén a közjót szolgálja. Ebben a műben is megmutatkozik Bacon tudományeszménye azáltal, hogy azonosítja a tudományos kutatást, mint az ideális társadalom alapját, felrajzolva egy olyan világ képét, ahol a tudományos fejlődés etikus keretek között, az emberiség javát szolgálva valósul meg és a tudósokat-feltalálókát általános megbecsülés övezi:

„Ami szokásainkat és szertartásainkat illeti, van két szép és tágas csarnokunk; az egyikben a ritka és kiváló találmányok mintáit helyeztük el, a másikban a neves

felfedezők szobrait. (...) sok közülünk való kiváló feltalálónak is van szobra, akik nagyszerű dolgokat és műveket találtak fel; ezeket a találmányokat leírni, mivel itt nincsenek előttünk, hosszadalmas lenne, és az ilyen dolgok megítélésében könnyű tévedni. Minden arra méltó dolog feltalálójának csakhamar szobrot állítunk és elég bőkezű tiszteletdíjat juttatunk” (Bacon, 2001, p. 22).

A műben felvázolt társadalmi berendezkedés, a tudósok számára széleskörű autonómiát biztosít; a technológiai fejlesztéseken túl, morális és közéleti kérdésekben is véleményt formálnak és a társadalom irányítóinak tanácsadóiként funkcionálnak. A tudomány és a hatalom viszonya ebben társadalomban szimbiotikus, a tudósok tevékenységének célja a fennálló rend fenntartása és javítása, ami által kivételezett helyzetüket nyerik. Érdekes ennél a szakasznál megemlíteni az újítás (innovation) és a feltalálás (invention) fogalmak közötti különbségtételt, amelyet már az előző fejezetben az innováció definíciói kapcsán említettem (Bacon az innovation kifejezést ebben a műben nem használja, az invention különféle szövegkontextusban viszont hét alkalommal szerepel). A „feltalálás” egy szellemi tevékenység, amely általában egy új, eredeti eszköz, gép, folyamat vagy anyag létrehozását eredményezi és gyakran egy technikai vagy tudományos áttörés eredménye; az innováció – ahogy az előző fejezetben megfogalmaztam – egy sokdimenziós folyamat, amely magában foglalja új vagy jelentősen továbbfejlesztett elméletek, termékek, szolgáltatások, technológiák, folyamatok, szervezeti formák vagy piaci megközelítések létrehozását, adaptációját, implementációját és diffúzióját – lényeg a gyakorlati alkalmazáson van.

A XXIV-es számú, „Az újításokról” (On innovations) címet viselő esszéjében Bacon az innováció fogalmát vizsgálva párhuzamot von a természetes születési folyamatok, valamint az új eszmék vagy gyakorlatok megjelenése között, azt sugallva, hogy az innovációk, akár csak az újszülött lények „*torzak eleinte*” (ill–shapen), elégtelenek-hibásnak tűnhetnek, mégis rendelkeznek azzal a potenciállal, hogy felülmúlják a meglévő gyakorlatokat. Az új ötletek megalkotóit a becsületes családok alapítóihoz hasonlítja, akik gyakran több erényt és méltóságot mutatnak, mint utódaik – ezáltal az új megközelítések úttörőinek egyedi és értékes hozzájárulását hangsúlyozza, azt sugallva, hogy az igazi innováció ritkán valósul meg pusztán utánpótlással. Óva int az elavult szokásokhoz való merev ragaszkodástól, azzal érvelve, hogy az ilyen konzervativizmus ugyanolyan bomlasztó lehet, mint a vakmerő innováció. Bacon a kiegyensúlyozott megközelítés mellett érvel, ahol a változtatásokat átgondoltan és fokozatosan hajtják végre, tükrözve magának az időnek a finom és egyenletes fejlődését: „*Úgy volna hát jó, ha az emberek az újításban magának az időnek a példáját követnék, amely bizony nagymértékben újít, de csendben és alig észrevehető fokozatokban*” (Bacon, 1987, p. 109).

Bacon óv az újdonságok önmagukért való hajszolásától, és különbséget tesz a valódi reformok között, amelyeket a fejlődés szükségessége vezérel, és a felszínes változások között, amelyeket a status quo iránti pusztán elégedetlenség motivál. Ez a megkülönböztetés döntő fontosságú az innovációval kapcsolatos etikai megfontolásai megértéséhez, mivel Bacon arra utal, hogy nem minden új ötlet vagy gyakorlat eredendően hasznos, az újítások egyesek számára kedvezőek, míg mások számára kedvezőtlenek lehetnek. Óvatosságra int, hogy az újdonságokat gyanakvással kezeljük, és a hagyományokra, a korábbi korok tudására támaszkodva keressük meg a helyes utat. igyekezve egyensúlyt teremteni az innováció és a stabilitás között.

A *Novum Organum* XXXI. szakaszában azonban a tudományra vonatkozóan egyértelműen egy változásorientált, az újításra nyitott szemléletet fogalmaz meg: „*Hiába remélünk jelentős haladást a tudományban attól, hogy a régit az újba töltögetjük és oltogatjuk; újra kell építenünk mindent az alaptól a tetőig, ha nem akarjuk örökké egy körben forogva, jelentéktelen és szinte említésre sem méltó haladással beérni*” (Bacon, 2001, p. 26). Kiemeli, hogy nem akarja lerombolni a már meglévő tudományokat és mesterségeket, hanem azok gazdagítására, továbbfejlesztésére törekszik, bízva abban, hogy a módszerek a jövőben tovább fejlődhetnek, különösen a gyakorlati felfedezések bővülésével párhuzamosan.

A *Novum Organumban* megfogalmazott gondolat, miszerint „újra kell építenünk mindent az alaptól a tetőig”, jól tükrözi Bacon elkötelezettségét az átfogó és mélyreható változások iránt, amelyek az emberiség javát szolgálják. Az innováció ebben az értelmezésben egy olyan folyamat, amely a tudományos módszerek és a társadalmi struktúrák újragondolását is magában foglalja, annak érdekében, hogy egy jobb, igazságosabb világot teremtsen. Bacon munkássága ezen a ponton nemcsak a múlt és jelen közötti átmenetet, hanem a jövő lehetőségeinek megalapozását is szolgálja, kijelölve a tudományos kutatás és az innováció etikai irányvonalait.

3.6 A társadalom változásorientáltsága Giambattista Vico filozófiájában

A Francis Bacon-t követő korszakban Giambattista Vico (1668–1744) az filozófus, akinek munkásságában az újítás – ha nem is explicit módon – szerepet kap. A nápolyi tudós „Az új tudomány” című úttörő művével forradalmasította a tudományos gondolkodást, a történelmi ciklusok, az emberi természet és a társadalmak fejlődésének feltárása révén új alapokra helyezve a civilizációk fejlődésének és a kultúra a világunk alakításában játszott szerepének megértését. A tapasztalat, a kulturális kontextus, a nyelv és a képzelet fontosságának hangsúlyozásával az emberi tudás alakításában, Vico megkérdőjelezte a karteziánus

racionalizmust, és utat nyitott az interdiszciplináris tudomány szemlélet előtt. Az emberi tudás és megértés dinamikus jellegét hangsúlyozta, amely a kulturális és társadalmi változásokkal együtt fejlődik. Ciklikus történelemszemléletében civilizációk fejlődésének három szakaszát különíti el – az istenek, a hősök és az emberek korát – amelyek közül minden korszaknak megvannak a maga sajátosságai a nyelv, a társadalmi szervezet és a gondolkodásmód tekintetében és mindezt az isteni gondviselés irányítja. A civilizációk felemelkednek, majd hanyatlani kezdenek és ha a társadalmak túlságosan elkorcsosulnak, visszatérnek egy primitívebb állapotba, és a ciklus újrakezdődik – így a társadalmak alá vannak vetve az állandó változás folyamatainak. Kiemeli az emberi tudásépítés narratív aspektusait, a világ megértésének módozatait bemutatva a mítoszokon és metaforákon keresztül történő értelmezésektől az absztrakt gondolkodásig. Vico felhívja a figyelmet, hogy az emberi tudás nem korlátozódhat pusztán az érzékelésre és a logikai dedukcióra; az emberi történelem, kultúra és nyelv megértése szintén elengedhetetlen, mivel ezek tanulmányozása révén az emberek jobban megérthetik saját természetüket és közösségeiket. Mivel a társadalmak emberi konstrukciók, azok kapcsán az ismereteink is bizonyosak, mivel az emberi megértés leginkább azon dolgokra terjedhet ki, amelyeket maga az ember alkotott vagy befolyásolt. Az 1710-ben megjelent, *De antiquissima Italorum sapientia* (Az egykori itáliaiak bölcsességéről) első fejezetében megfogalmazza Vico ezt az alaptételét:

*„Atque ex his, quae sunt hactenus dissertata, omnino colligere licet, veri criterium ac regulam ipsum esse fecisse: ac proinde nostra clara ac distincta mentis idea, nedum ceterum verorum, sed mentis ipsius criterium esse non possit: quia dum se mens cognoscit non facit; et quia non facit, nescit genus, seu modum, quo se cognoscit.”*¹⁶ (Vico, 1710, p. 26).

A Vico által meghatározott *verum et factum convertuntur* axióma alapján az emberi tudás alapvetően az ember által létrehozott dolgok megértéséből származik, és az igazság (verum) és a megalkotott (factum) felcserélhető. Azáltal, hogy az ember által alkotott dolgok megismerhetők számunkra, a társadalmi valóság elemei is feltárhatók, meghatározhatók a fejlődési elvei, amelyek az embert erkölcsi és politikai lényé formálják. A fejlődésben fontos szerepe van az isteni Gondviselésnek (Vico a filozófiájában megtartja a metafizikai alapokat), a civilizáció fokozatosan halad a fejlettebb állapotok felé, bár a ciklikus szemlélet szerint időszakosan visszatér a kezdeti állapotokhoz. Az emberi társadalmak fejlődésével maga a

¹⁶ „Az eddig tárgyaltak alapján teljesen világosan következtethetünk arra, hogy az igazság kritériuma és mércéje maga a létrehozás: ezért a mi világos és határozott mentális képünk nem lehet sem az egyéb igazságok, sem önmagunk mércéje, mert amikor az elme önmagát felismeri, nem alkot; és mivel nem alkot, nem tudja, hogy milyen módon ismeri fel önmagát.” (saját fordítás)

társadalmi rend és a politikai-hatalmi struktúrák is folyamatosan változnak a különböző társadalmi osztályok közötti feszültségek és érdekellentétek miatt: „...*a plebejusok mindig meg akarják változtatni a meglevő állapotokat, s mindig változtatnak is rajtuk; a nemesek viszont mindig fenn akarják tartani őket*” (Vico, 1953, p. 385).

Az idézett mondat, amelyben Vico a plebejusokat és a nemeseket említi, a társadalmi dinamizmus egy fontos aspektusát ragadja meg: az állandó konfliktust a változásra törekvő alsóbb osztályok (plebejusok) és a fennálló rendet megőrizni kívánó felsőbb osztályok (nemesek) között. Machiavellihez hasonlóan Vico is elismeri, hogy a konfliktusok mind kreatív, mind destruktív erővel bírhatnak, a társadalmi dinamika és a történelmi fejlődés motorja gyakran az ellentétes érdekű csoportok közötti konfliktusokban rejlik, az alsóbb és felsőbb osztályok közötti feszültségek és ellentétek olyan helyzeteket teremtenek, amelyek változásra kényszerítik a társadalmat. Ez a gondolat összhangban van Vico általános elméletével, amely szerint a történelem ciklikus, és a társadalmak folyamatosan változnak és fejlődnek az emberek cselekedetei és törekvései révén.

Giambattista Vico filozófiája jelentős mértékben hozzájárul az innováció történeti és elméleti megértéséhez. A *verum et factum convertuntur* elv, a társadalmi fejlődés hármas ciklusa, valamint a nyelv, kultúra és tapasztalat hangsúlyozása mind olyan aspektusok, amelyek új perspektívát nyitnak az innovációs folyamatok értelmezésében. Vico gondolatai rávilágítanak arra, hogy az innováció nem pusztán technikai vagy tudományos kérdés, hanem mélyen gyökerezik a társadalmi dinamikában, az emberi gondolkodás fejlődésében és a kulturális kontextusban. Az isteni inspiráció és az emberi cselekvés kettőssége, a mítoszoktól az absztrakt gondolkodásig ívelő tudásépítés, valamint a társadalmi feszültségek, mint a változás motorjai mind olyan elemek, amelyek Vico filozófiájában az innováció komplex, többdimenziós természetét tükrözik. Ez a megközelítés nemcsak történeti szempontból értékes, de a jelenkori innovációs folyamatok mélyebb megértéséhez is hozzájárulhat, hangsúlyozva az emberi kreativitás, a társadalmi kontextus és a kulturális fejlődés elválaszthatatlan kapcsolatát az újítások létrejöttében és terjedésében.

3.7 Marx és a technológiai forradalom dialektikája: az innováció ígérete és paradoxonjai

Vico *verum ipsum factum* elve azt állítja, hogy csak azt érthetjük meg teljesen, amit létrehozunk vagy véghezviszünk, mivel cselekvésünk adja a tudás alapját. Ez az elv megkérdőjelezi a karteziánus megközelítést, amely a deduktív módszerre épülő fizikát tartotta a legmegbízhatóbb tudásformának és többre értékelte a történelmi és nyelvi ismeretekkel szemben. Vico mellett

érvelt, hogy az ember, mint aktív formálója annak a világnak, amelyet társadalmi, politikai és kulturális tevékenységei révén létrehoz, jobban képes megérteni ezáltal magát a világot is, amihez értékes eszközt kínál a történelem és a hagyományok ismerete. Vico koncepciója, amely szerint az emberi ágenciát helyezi előtérbe a történelem értelmezésében, a 19. században Karl Marx (1818-1883) gondolkodásában köszön vissza – habár ezen túl a kapcsolatuk nem kimutatható (Martins, 2022, pp. 26-32).

Marx gondolkodásában – csakúgy, mint Vico esetében – az innováció fogalma közvetlen módon nem kap kiemelt figyelmet, de implicit módon jelen van a műveiben: részletesen vizsgálja a technológiai fejlődés társadalmi hatásait, a termelőerők fejlődését, és a termelési módok átalakulását, amelyek mind összekapcsolhatók az innováció fogalmával. Guido Frison rámutat, hogy Marx munkásságban értékes meglátásokat találunk az innováció és a gazdasági fejlődés folyamatairól. Marx részletesen elemezte a technológiai változások és az innováció terjedésének folyamatát, különös tekintettel a régi és új technológiák közötti versenyre. Az innovációt endogén jelenségként értelmezte, amely a munkafolyamatok módosulásából ered, azaz az innováció a termelési folyamat szerves részeként jelenik meg, nem külső tényezők hatására. A változások a munkafolyamat belső dinamikájából fakadnak, például a munkamegosztás fejlődéséből vagy a termelési eszközök átalakulásából. Szerinte az innováció szorosan összefügg a termelési módok változásával, a kézműves termeléstől a manufaktúrán át a modern gyáripár kialakulásáig (Frison, 1988, pp. 309-324).

A modern ipar felbomlasztja a hagyományos munkamegosztást (és egyben a régi családi viszonyokat), de emellett forradalmasítja a termelési folyamatokat és a munkások szerepeit és funkcióit. Ez a folyamat egyrészt negatív hatásokkal jár: bizonytalanságot, a munkások kizsákmányolását és társadalmi feszültségeket eredményez, munkaerő specializálódása pedig a munkások képességeinek fejlődésében szab korlátokat. Másrészt azonban Marx szerint ez a folyamat szükségszerű a fejlődéshez, ezek a változások kényszerítik ki a munkások adaptációs képességét, ami hosszú távon pozitív változást hozhat:

„A modern ipar egy termelési folyamat meglevő formáját soha nem tekinti és kezeli véglegesnek. Technikai bázisa ezért forradalmi, míg minden korábbi termelési módé lényegileg konzervatív volt. Gépi berendezés, vegyi folyamatok és más módszerek révén a modern ipar állandóan forradalmasítja a termelés technikai alapjával együtt a munkások funkcióit és a munkafolyamat társadalmi kombinációit. Ezzel éppoly állandóan forradalmasítja a munkának a társadalmon belüli megosztását, és szakadatlanul töketömegeket és munkástömegeket dob egyik termelési ágból a másikba. A nagyipar természete ezért feltételezi a munka változását, a funkció folyékonyágát, a munkás mindenoldalú mozgékonyágát” (Marx, 1955, p. 454).

Marx A politikai gazdaságtan bírálatának alapvonalai-ban (Marx, 1972, pp. 159-170) kifejti, hogy a technológiai fejlődés és az innováció alapvetően megváltoztatja a munka természetét és a gazdaság működését. Szerinte a valódi gazdagság forrása már nem csupán a közvetlen munkaidő, hanem inkább a tudomány és technológia alkalmazása a termelésben. Ez az átalakulás azt eredményezi, hogy a munkás szerepe átalakul: már nem közvetlenül a termelés fő ágense, hanem inkább felügyelője és szabályozója a termelési folyamatnak.

A tőkés osztály folyamatosan forradalmasítja a termelőeszközöket, ezáltal a termelési és a társadalmi viszonyokat is. Marx rávilágít arra, hogy a munkásosztály által létrehozott technológiai fejlődés végül magát a munkásosztályt fenyegeti, paradox módon a gép, amely a munkásosztály erőfeszítéseinek eredménye, gyakran a munkás ellen fordul, kiszorítva őt a munkahelyéről, illetve fokozva kizsákmányolását. A marxi elmélet kritikus a technológiával szemben, annak elidegenítő hatását hangsúlyozza – másfelől a technológiai fejlődés jövőformáló lehetőségeit is látja, amennyiben a kapitalizmus és a technológiai fejlődés ellentmondásba kerülnek egymással, ami a rendszer válságához vezet. Marx szerint a technológiai fejlődés és a tudományos ismeretek alkalmazása egyre fontosabbá válik a gazdasági növekedés szempontjából, miközben megkérdőjelezi a hagyományos munkaérték-elméletet. Marx szerint ez a folyamat végül a kapitalista rendszer összeomlásához és egy új társadalmi rend kialakulásához vezethet. Amy Wendling a Karl Marx on technology and alienation című könyvében kritizálja Marx hajthatatlanságát a forradalmi változással kapcsolatban, ami figyelmen kívül hagyja a fokozatos fejlődés lehetőségét és a munkások technológiai ágenciáját (Wendling, 2009, pp. 190-192).

Marx bár kritikusan szemléli a kapitalista rendszer keretein belül végbemenő változásokat, felismeri a fejlődés elkerülhetlenségét és potenciális pozitív hatásait is. Ez a dialektikus szemlélet, amely a fejlődést egyszerre tekinti romboló és építő folyamatnak, átvezet minket Friedrich Nietzsche gondolatvilágához. A Schumpeter kapcsán korábban már idézett *teremtő rombolás* elmélete szerint (Schumpeter, 1980, pp. 111-112) az innovációs folyamat lényege, hogy új kombinációk jönnek létre, amelyek a régi rendszerek helyébe lépve, versenyre kelnek a megszokott struktúrákkal, fokozatosan elszívják erőforrásaikat, és végül kiszorítják őket, ami egyesek felemelkedését, mások hanyatlását eredményezi. A Marx-i gondolkodásban ez a fent már tárgyalt állandó technológiai és gazdasági újításokat jelenti a kapitalizmusban, amelyek gyakran megsemmisítik a régi struktúrákat és technológiákat, új és hatékonyabb termelési módszereket vezetve be.

3.8 Nietzsche hatása a teremtő rombolás elméletére

Míg Marx a társadalmi-gazdasági struktúrákra összpontosított, Nietzsche (1844-1900) az egyén és a kultúra szintjén vizsgálta a fejlődés, a változás és a rombolás kérdéseit. Nietzsche filozófiájában szintén megjelenik az a gondolat, hogy a fejlődés nem mindig egyenes vonalú és problémamentes folyamat, hanem gyakran kríziseken és radikális átértékeléseken keresztül valósul meg, amelynek részét képezi a meglévő struktúrák lerombolása.

Általában Schumpeternek tulajdonítják a teremtő rombolás gondolatának a közgazdaságtanba való bevezetését, Hugo Reinert és Erik S. Reinert szerint azonban annak a gyökerei sokkal mélyebbre nyúlnak, eredeténél hinduizmustól a reneszánsz filozófián át Friedrich Nietzsche-ig tartó gondolati szálat kell végigkövetnünk. Tanulmányukban árnyalt értelmezést adnak Nietzsche az Így szólott Zarathustra-ban artikulálódó kreatív pusztításról alkotott perspektívájáról, a teremtés, a pusztítás és a megőrzés fogalmai közötti feszültségként írva le azt (Reinert & Reinert, 2000). A Zarathustra-ban Nietzsche két ellentétes alakot mutat be: a nemest, aki mind az alkotást, mind a pusztítást megtestesíti, illetve a jókat és igazakat, akik a megőrzést, a stagnálást és a hanyatlást képviselik. Ezt a kettősséget a hatalom akarása fogalmán keresztül magyarázza, amely a nemest alkotásra ösztönzi, míg a jók és igazak híján vannak ennek az alkotói késztetésnek, és ehelyett a fennálló rend fenntartására törekszenek: „*A nemes útjában áll bizony még a jóknak is: ha pedig jónak mondják, csakis azért, hogy könnyebb legyen félreállítaniuk. Újat akar teremteni a nemes, és új erényt. A régit akarja a jó, és hogy a régi megmaradjon*” (Nietzsche, 2000, p. 55, 1025. szakasz). Az újraértékelés nem pusztán destruktív tevékenység, hanem egy aktív, teremtő folyamat egyben, ami az emberi életet és kultúrát magasabb szintre emeli, egyfajta szellemi evolúció eszköze, amelynek a célja, hogy a legalkalmasabb és legéletképesebb eszmék maradjanak fenn a filozófiai és kulturális szelekció folyamatában. Az Adalék a morál genealógiájához -ban Nietzsche így fogalmaz: „*Ha szentélyt akarunk építeni, előzőleg le kell rombolnunk egy másik szentélyt: ez törvény, mulassatok nekem olyan esetet, ahol nem e törvény érvényesült!*” (Nietzsche, 1996, p. 24). Reinert hangsúlyozza, hogy a hatalom akarásának természete dinamikus, vagy emelkedik, vagy hanyatlik, ami állandó önmeghaladást tesz szükségessé. Nietzsche azt állítja, hogy a régi értékek és hiedelmek lerombolása szükséges az új, magasabb rendű eszmények és gondolkodásmódok megteremtéséhez - ez a kreatív rombolás folyamata központi szerepet játszik az önmeghaladás filozófiájában és az emberi potenciál folyamatos megújulásában. Zarathustra alakján keresztül mutatkozik be: „*Bizony mondom néktek: még káosz kell legyen abban, aki táncoló csillagnak akar életet adni*” (Nietzsche, 2000, p. 25, 235 szakasz).

Két olasz közgazdász, Enrico Santarelli és Enzo Pesciarelli (1990) tanulmányukban Joseph Schumpeter filozófiai és gazdasági elméleteit elemzik, kiemelve a vállalkozói szellemről alkotott koncepcióját és annak kapcsolatát Friedrich Nietzsche gondolataival. Felhívják a figyelmet, hogy Schumpeter vállalkozó-felfogása összhangban áll Nietzsche emberfeletti ember koncepciójával, mindkét fogalom olyan egyéneket ír le, akik túllépnek a konvencionális normákon a változás és innováció előmozdítása érdekében. Ezek az alakok szemben állnak a *tömeggel* vagy *nyájjal*, akik a bevett utakat követik és ellenállnak a változásnak. Ahogy Nietzsche fogalmaz: „*A világ legjobb dolgai is értéktelenek, ha nincs, aki színre vigye őket; az ilyen színrevivő fajtát hívják úgy: nagy emberek. A sokaság érzéketlen a nagyság iránt, ami - a teremtő. Van viszont érzéke a nagy dolgok összes színrevivőjéhez és színészéhez. Új értékek kitalálói körül forog a világ - forog láthatatlanul*” (Nietzsche, 2000, p. 65, 1245. szakasz). Schumpeter vállalkozója nem csupán egy vállalkozás menedzsere, hanem a dinamikus változás ügynöke, aki új kombinációkat és innovációkat keres, amelyek megzavarják a statikus gazdasági feltételeket. Schumpeter gazdasági fejlődésről alkotott nézete ciklikus (akárcsak Vico, Marx vagy Nietzsche esetében), ahol az innovációk olyan változásokat indítanak el, amelyek új egyensúlyi állapotokhoz vezetnek a zavarok időszakai után – ezt a ciklikus jelleget az innovációk időzítése és szektorok közötti eloszlása befolyásolja. Schumpeter arra utal, hogy míg a legtöbb ember alkalmazkodik környezetéhez, az igazi változás azoktól a ritka egyénektől származik, akik ösztön, akaratérő és innováció kombinációjával kitörnek koruk korlátai közül, hogy valami újat hozzanak létre – mintha Zarathustra alakjáról beszélne: „*Látnok, akaratnak és teremtésnek embere, maga a jövő és híd a jövőbe*” (Nietzsche, 2000, p. 171, 1774. szakasz). Ez a vízió kihívást jelent a klasszikus közgazdaságtan statikus, egyensúly-központú modelljeivel szemben, ehelyett egy dinamikus, ciklikus fejlődési folyamatot javasol, amelyet az emberi cselekvés hajt. Azt az intellektuális miliőt vizsgálva, amely Schumpeter gondolatait formálta, mindenképpen figyelemmel kell lennünk Nietzsche hatására, amely segített Schumpetert egy olyan történelem- és gazdaságfelfogás kialakításában, amely hangsúlyozta az energikus egyének szerepét a társadalmi evolúció előmozdításában, gyakran kiszámíthatatlan módokon (Santarelli & Pesciarelli, 1990, pp. 689-694).

3.9 Összefoglalás

A fejezetben végigkövettem az innováció fogalmának megjelenését és fejlődését – némi kitérőt téve – a nyugati szellemtörténetben, az ókori Görögországtól a 19. századig. Az elemzésből kiderül, hogy az innováció értelmezése és a fogalom használata jelentős átalakuláson ment

keresztül, amelyet a különböző filozófiai iskolák és az egyes korszakok társadalmi-kulturális kontextusai alakítottak.

Az ókori görög filozófiában az innovációhoz való viszonyulás ambivalens volt. Miközben a klasszikus korszakban számos területen figyelemre méltó fejlődésnek lehettünk tanúi, az olyan filozófusok, mint Platón és Arisztotelész óvatosak voltak az innovációnak a társadalmi stabilitásra és a hagyományos értékekre gyakorolt esetleges bomlasztó hatásaival kapcsolatban. Hangsúlyozták a régi és az új közötti egyensúly fenntartásának fontosságát, valamint azt, hogy a változások bevezetése előtt alapos megfontolásra van szükség.

A középkori iszlám filozófia, a vallási tanításokat igyekezett összeegyeztetni a racionális vizsgálódással és a gyakorlati megfontolásokkal. Különbséget tettek a hasznos és káros újítások között, hangsúlyozva annak fontosságát, hogy az új eszméket az iszlám elveivel és a közjával hangolják össze. Ez a megközelítés megalapozta a későbbi keresztény skolasztikát, amelynek egyik legfontosabb képviselője Aquinói Tamás volt, aki szintén a hit és az ész kapcsolatával foglalkozott – esetében a törvények kapcsán találkozhattunk az újítás kérdésének tárgyalásával.

A reneszánsz fordulópontot jelentett az innováció értelmezésében, amikor olyan filozófusok, mint Machiavelli és Bacon egy proaktívabb és pragmatikusabb megközelítés mellett érveltek. Machiavelli az innovációban a hatalom fenntartásának és a változó körülményekhez való alkalmazkodás eszközét látta, míg Bacon az empirikus megfigyelés és kísérletezés fontosságát hangsúlyozta a tudás fejlődésében és szerepét a társadalmi állapotok és az emberi jólét javításában. Mindkét gondolkodó azonban felismerte az ellenőrizetlen innovációval járó lehetséges kockázatokat és etikai dilemmákat is – ebben fontos előfutárai kortárs gondolkodóknak.

A 18. és 19. században az innovációról folytatott diskurzusban további elmozdulás következett be, amikor olyan gondolkodók, mint Vico, Marx és Nietzsche, feltárták az innováció összetett és gyakran paradox természetét. Vico ciklikus történelemszemlélete és az emberi cselekvés szerepének hangsúlyozása a társadalmi valóság alakításában implicit hatást gyakorolt az innováció későbbi értelmezéseire. Marxnak a technológiai fejlődés és a társadalmi változások közötti kapcsolatáról szóló elemzése rávilágított az innováció dialektikus természetére, amely egyszerre teremtő és romboló erő, míg a teremtő rombolás koncepciója Nietzsche-nél és a kivételes egyének szerepének hangsúlyozása a változások előmozdításában további hozzájárulást jelent a modern innovációértelmezés megszületéséhez.

A fejezet célja az volt, hogy az nyugati filozófiatörténet időfolyamában vizsgálva az innováció fogalmát, feltárjam annak sokrétű és gyakran ellentmondásos természetét. Ez a kettősség mindvégig megmutatkozik: miközben az innovációt a haladás hajtóerejeként és a változó körülményekhez való alkalmazkodás eszközeként ismerték el minden korban, óvatossággal és a társadalmi stabilitásra és a hagyományos értékekre gyakorolt lehetséges bomlasztó hatásaival kapcsolatos aggodalmakkal is találkozunk szinte minden szerzőnél. A régi és az új, a hagyomány és az újítás közötti feszültség visszatérő témaként jelenik meg az innovációról szóló filozófiai diskurzusban és mindvégig érződik a kihívás, amivel a gondolkodók küzdenek, hogy egyensúlyt találjanak a múlt bölcsességének megőrzése és a jövő lehetőségeinek felvállalása között. Szinte minden szerző felismeri az újítás etikai dimenzióit, és az egyensúly megtalálásának fontosságát a közjót szolgáló innováció és a stabilitást veszélyeztető öncélú újítás között.

Ami ebből a fejezetből egyértelműen kiviláglik, az az innováció tágabb társadalmi, kulturális és politikai összefüggésekkel való szoros kapcsolata. Az egyes történelmi korszakoktól és társadalmi változásoktól függően változott az innováció megértése és megítélése, és ez fordítva is igaz: az innováció természetével kapcsolatos filozófiai felismerések is alakították azt a módot, ahogyan a társadalmak megközelítették és kezelték a változásokat.

A fejezet szándékoltan adós marad az innováció vizsgálatával nyugati keresztény vallási kontextusban – annak ellenére, hogy az iszlám filozófia kapcsán tárgyalja a vallási újítás kérdéskörét. Ennek több oka van: egyfelől a téma kiterjedtsége és jelentősége a nyugati kultúrtörténetben olyan mértékű, hogy a vallási dimenzió részletes elemzése túlmutat ezen dolgozat fókuszán. Másrészt, a téma a keresztény hagyományban olyan példákon keresztül mutatható be, amelyek referenciapontjai kívül esnek mind a fejezet céljain, mind a témám tudományos relevanciáján. Ebben a tárgykörben – a már többször citált – Benoit Godin *The English Reformation and the Invention of Innovation* (Godin, 2020b) című tanulmánya nyújt elmélyült elemzést.

Elmondható, hogy az innovációnak a nyugati filozófia szemüvegén keresztül történő vizsgálata értékes tanulságokkal szolgál a témával kapcsolatos kortárs viták számára. A következő fejezetekben az innováció felelősségetikai megközelítésének elemzésén keresztül, ennek a reflexiónak a kiszélesítésére teszek kísérletet.

IV. Az innováció az erkölcs és a technológiai fejlődés ütközőzónájában

4.1 A tudomány új horizontja: az innovációvezérelt kutatás lehetőségei és kihívásai

Ahogy az előző fejezetben végigkövethettük, a történelem során a felfedezés és a fejlődés vágya volt a tudományos kutatások elsődleges motivációja, és a cél „*hogy új találmányokkal és javakkal gazdagítsák az emberi életet*” (Bacon, 1995, LXXXI. szakasz, s.p.). Ezzel szemben egyre nagyobb aggodalomra ad okot napjainkban, hogy mind az akadémiai, mind a vállalati szektorban megfigyelhető egyfajta *innovációs kényszer*. A meglévő tevékenységek és eredmények továbbfejlesztésével és javításával szemben előnyben részesülnek az új felfedezések, amelyeket *ösztönzők* bevezetésével próbálnak előmozdítani - következésképpen jelentős nyomás nehezedik a fejlesztőkre, újabb és újabb innovatív kutatások gyors elvégzésére sarkalva őket (Sarewitz, 2016). Fontos megjegyezni azonban, hogy az innovációnak ez a sajátos, erőltetett formája az elmúlt évtizedekben jelentős előrelépéseket is hozott, amelyek a társadalmi előnyei számtalanok, így a kép is sokkal árnyaltabb.

Számos kutató szerint a tudományos felfedezések és a technológiai áttörések alkalmazása jelentős előnyökhöz vezetett az emberiség számára, amelyek javították az életminőséget és pozitív társadalmi változásokat eredményeztek (Roco & Bainbridge, 2003). A tudományos eredmények közvetlen összefüggést mutatnak az orvostudomány, a mezőgazdaság, az energia, a közlekedés és a számítástechnika terén elért fejlődéssel. Az orvostudományi kutatások számos új gyógyszer és betegségterápia létrejöttét eredményezték, amelyek forradalmi változásokat hoztak az egészségügy területén. A mindinkább elterjedő újszerű energiatermelési technológiák, például a nap- vagy szélenergia bevezetése is radikális változásokat hozott a globális közösség geopolitikai és gazdasági horizontján. Az olyan tudományágak, mint a genetika, a neurológia, a nanotechnológia és a mesterséges intelligencia terén elért előrelépések további jelentős áttörések lehetőségét kínálják, ugyanakkor új etikai dilemmákat is felvetnek (Jasanoff, 2021). Fontos elismerni a tudományos felfedezések jelentős előnyeit, ugyanakkor elengedhetetlen a felelősségteljes tervezés is, az előre nem látható következmények minimalizálása és az előnyök igazságos elosztásának biztosítása érdekében

A kritikusok azt állítják, hogy az innovációhoz elengedhetetlen tudás megszerzését veszélyeztetheti a modern tudományosságban eluralkodó rövid távú gondolkodás, amelyet a gyors eredményekre törekvés motivál. Ez olyan módszerekhez vezethet (Brainard, 2018), amelyek alááshatják a kutatások hitelességét, így például terjed az etikailag kifogásolható

kutatási eljárások alkalmazása, a nem megfelelően megválasztott módszertan alkalmazása, és a téves vagy túlzó következtetések levonása. Az ellenőrizetlen innováció negatív hatásai a környezeti fenntarthatóságra is egyre nyilvánvalóbbak – a mesterséges intelligenciával kapcsolatos, vagy egyes biotechnológiai és a nanotechnológiai fejlesztések jól példázzák ezt a problémát. A fenntartható és ellenőrzött innovációs tevékenységhez elengedhetetlen az etikátlan motivációk kezelése és annak újragondolása, hogy a szervezetek miként értékelik és ösztönzik a tudományos kutatómunkát, milyen morális ellensúlyokat építenek be a kutatási folyamatokba.

Nicole N. Morris egy paradigmaváltás szükségessége mellett érvel, az eddiginél elszámoltathatóbb keretek kialakítását javasolva. Az innováció átfogó etikai keretrendszerének hiánya folyamatos kihívásokat jelent mind a fejlesztői oldal, mind a társadalom számára – állítja. Egyfelől a kutatók egyértelmű útmutatások, meghatározott normák nélkül kényszerülnek az etikai dilemmák között kiigazodni, másfelől a befektetői oldal számára a nem tervezett potenciális károk kalkulálhatatlan kockázatként jelennek meg, visszavetve a finanszírozási szándékot. Mindez kiegészül a potenciális csalásokon, hamisított információkon alapuló fejlesztési tevékenységek okozta károkkal (Morris, 2024, p. 41), amelyek szabályozására és visszaszorítására ajánl Morris etikai és jogi eszközöket, egyfelől a visszaélések bejelentését ösztönző, másfelől egy retributív szemléletet népszerűsítve.

A piacósítható új eredmények létrehozására irányuló nyomás megkérdőjelezhető eljárások alkalmazásához vezethet, az úttörő felfedezésekért járó jelentős jutalmak és a fennálló intenzív verseny nagy nyomást helyez a kutatókra. A tudományos közösség egyre inkább üzletiesedő szemlélete a kutatókat arra ösztönzi, hogy a kollaboratív és kooperatív megközelítések helyett, a saját eredményeikért járó elismerést helyezték előtérbe. Ez a megközelítés a tudás fragmentálódásához vezet, izolált tudásbázisok létrejöttét hozza el, amellyel, hogy a nem kellő körültekintéssel, gyorsan lefolytatott kutatások általában alacsonyabb színvonalúak, így a kortárs tudományos közösségben kialakult erős versenykörnyezet, a tudomány eredendő célját jelentő tudásfejlesztés rovására megy. Elengedhetlenné válik bizonyos reformok végrehajtása a fenntartható kutatási kultúra előmozdításához (Szilágyi, 2024a, pp. 142-147).

4.2 A felelősség és az ágencia kérdései az innovációetikában

A 21. század második évtizedére a technológiai innovációk olyan katalizátorként jelentek meg azokban a társadalmakban, amelyek számára biztosított a hozzáférés ezen fejlesztésekhez,

amelyek korábban példa nélkül álló lehetőségek és beláthatatlan kockázatok egyidejű jelenlétét hozták el. Ezek a forradalmi technológiai változások nem csupán az innovációs ökoszisztémák robbanásszerű kiépülését és a már említett *innovációs kényszer* globális kialakulását jelentik, hanem kitörölhetetlenül átalakították az etikai gondolkodás körvonalait is. A gyors technológiai fejlődés eme turbulens közegében az érintett szervezetek világszerte kettős feladattal küzdenek: kiaknázni ezen innovációk kapacitásait és transzformatív potenciálját, ugyanakkor megoldásokat találni a bennük rejlő sebezhetőség kiküszöbölésére és olyan etikai kereteket meghatározni, amelyek biztosítják az emberi faj fennmaradását, jólétét és a társadalmi egyenlőtlenségek csökkenését.

Az előző fejezetben áttekintettem, hogy az innováció, az újítás (változás-változtatás) milyen kontextusban bukkant fel a filozófiatörténetben a 19. század végével bezárólag. Példáim többségében az innovációban rejlő kockázat nem technológiai oldalról jelenik meg, főként a társadalmi-politikai rendet fenyegetik az újítás következményei. Az elmúlt közel egy évszázad számos változást hozott, és az innováció értelmezési tartománya jelentős mértékben bővült, azonban tárgyalásában mindinkább egy technológiai fókusz érvényesül és egy kockázatközpontú megközelítés kerül előtérbe (ez például az ötödik fejezetben elemzett szabályozási dokumentumok esetében lesz hangsúlyos). Az innováció korábban általam megadott definíciójára visszautalva: azáltal, hogy az innováció ontogenetikus természetű, *„mélyreható változásokat hozhat létre a társadalmi-gazdasági rendszerekben, amely során új létezési módok és rendszerek jönnek létre, gyakran előre nem látható módon”* – magukban hordozva nem pusztán etikai dilemmák sorát, hanem a humán egzisztencia egészére fenyegetést jelentő események potenciális bekövetkeztét.

A technológiai fejlődés kockázatainak kérdése nem először merül fel a nyugati filozófiai hagyományban, számos szerző foglalkozik a kérdéssel. Így például Martin Heidegger (1977), aki szerint a modern technológia egy olyan szemléletmódot képvisel, amely a természetet puszta erőforrásként kezeli, amit ki lehet aknázni és uralni – ezzel radikálisan átalakítva az ember és a természet viszonyát. A technológia felemelkedése veszélyezteti az emberi létezés ontológiai alapjait, mert az emberek gondolkodásmódját és cselekvését egyre inkább a technológiai keretek határozzák meg, elveszítve a létezés mélyebb dimenzióival való kapcsolatot. Heidegger szerint a technológia alapvetően ontológiai kihívás, rákényszerítve az embert, hogy szembesüljön saját létezésének feltételeivel és újra megtalálja helyét a világban, egy funkcionális látásmódot kényszerítve rá. *„Az ember fenyegetettségét nem a technika helyell-közzel halált hozó gépei és készülékei hozták meg. A tulajdonképpeni fenyegetettség már ott van*

az ember lényegében” (Heidegger, 2004, p. 127) – írja. A technika kettős természetű, egyfelől valóságtorzító hatása van, másfelől, ha túltekintünk a technikán, mint eszközön, a megértésünk feltárja korlátait és kezelésének módozatait, segítve egy új viszony kialakítását a technikai világhoz.

A bevezetésben már idézett Mitcham (1994) felhívja a figyelmet Jacques Ellul munkásságára, aki a technológia totalitárius jellegét és autonómiáját elemzi részletesen a modern társadalomban. Az autonómia azt jelenti, hogy a technológia önálló entitássá válik, amelynek fejlődése kilép az emberi ellenőrzés és befolyás alól, aminek az eredményeképpen a társadalom egyre jobban alkalmazkodik az általa szabott követelményekhez. A hagyományos ember-gép kapcsolat megváltozott, ma már nem az ember irányítja a technológiát, hanem az ember maga is a technikai rendszer része, az emberi tevékenységek technikai rendszerek szabályai és igényei szerint alakulnak, így az ember valójában a technológiai fejlődésnek alávetett. A technológia egy olyan paradox helyzetet teremt, amelyben az ember növekvő természet feletti hatalma egyúttal együtt jár az autonómiájának a csökkenésével, a technológia a gondolkodását és a cselekvési lehetőségeit is meghatározza, habár meg van győződve, hogy ő irányítja azt: *„Választhat, de egy technikai folyamat által kialakított opciórendszerben. Irányt nyithat, de csak a technikai feltételeknek megfelelően. Egy pillanatra sem léphet ki a rendszerből”* (Ellul, 2004, p. 21). Ellul rávilágít, hogy a technológia nem pusztán eszközök és gépek összessége, hanem egy önállóan fejlődő rendszer, amely meghatározó módon formálja a modern társadalmak struktúráját és életmódját. Szerinte a technológia nem semleges, a technológiai fejlődés a saját belső logikája mentén halad előre, gyakran figyelmen kívül hagyva az emberi értékeket és döntéseket (Ellul, 1964).

Az előzőhöz hasonlóan Peter-Paul Verbeek is amellett foglal állást, hogy az ember és a technológia viszonya komplex interakciók sorozata, de egyben túl is mutat azokon. (Verbeek, 2015). A *Beyond Interaction* címet viselő tanulmányában három megközelítést vázol fel az ember és technológia viszonyának. Az egyikben a technológia egyfajta kiterjesztése (extension) az embernek, eszközként jelenik meg és segíti az emberi tevékenységet – a státusza semleges. A második, egy általa dialektikusnak nevezett viszony, amelyben a technológia és az ember között ellentét feszül, a technológia, mint egy elidegenítő erő jelenik meg (a marxi értelmezést hozza példaként). Végül a harmadik – hibrid – megközelítésben az ember és a technológia nem egymástól elhatárolt, utóbbi a része az emberi természetnek, illetve kölcsönösen visszahatnak egymásra: *„Our perceptions and experiences, our actions and ways of living, all these elements*

*of human existence take shape in close interaction with technologies*¹⁷ (Verbeek, 2015, s.p.). A funkcionalitásától eltávolodva, egyfajta közvetítő (mediation) szereplőként tekint a technológiára, amely segítségünkre van a minket körülvevő valóság feltárásában, a világ megismerésében, továbbá erkölcsi és metafizikai kereteink kialakításában. A technológia olyannyira a létünk része, hogy „*nélküle nem lehetünk emberek*”, és ahelyett, hogy a technológia felmérhetetlen hatásainak a kiküszöbölésére törekednénk – érvel Verbeek – ki kell hoznunk belőle a legjobbat, tudatában ennek a közvetítő hatásnak, körültekintően kell szabályoznunk és alkalmaznunk. Verbeek egy általa posztfenomenológiai nevezet megközelítést javasol, amelyben elfogadjuk azt, hogy a világot számunkra a technológia közvetíti és aktívan segítségünkre van abban, hogy kapcsolatba lépjünk vele (szemben a heideggeri koncepcióval, ahol a technológia elsősorban a világ értelmezésének módja). Az emberek az eszközök, a tárgyak segítségével érzékelnek és cselekszenek, így kerül az ember és a világ között egy közvetítő kapcsolatba a technológia, ami egyben alakítja is viszonyainkat és meghatározza hogyan bontakozik ki abban a létezésük (existence). A technológia tervezése inentől nem pusztán egy belső szakmaspecifikus folyamat, hanem erkölcsi kérdés is (Verbeek, 2005, p. 235), „*a technológia ugyanis mindig morális*” – mondja egy interjúban (Fatum, 2017). Nem elégséges önmagában a technológia hatását vizsgálni, a dinamikákat is elemezni kell, amelyek az ember-technológia viszonylatában működnek, azonosítva a felmerülő etikai kérdéseket és ez alapján kialakítani egy gyakorlati értékelést szolgáló keretrendszert, ami az erkölcsileg elfogadható tervezés eszköze is egyben (Verbeek, 2016, p. 196). Ez a szemlélet lényegében kiemeli a technológiát az embernek alárendelt szerepéből és a morális közösség aktív részévé teszi azt, mivel a közvetítés révén az ember általa nem csak a világot értelmezi, hanem a cselekedeteire is döntő befolyással bír. A felelősség és az ágencia így egy ember és egy nem humán entitás között oszlik meg, ami a gyökeresen új alapokra helyezi az etikát (Verbeek 2011, pp. 153-154).

Martin Sand szerint egyáltalán nem válik érthetőbbé a technológiai fejlődés kapcsán felmerülő kérdéseket tárgyaló etika azáltal, hogy az ember és a technológia ilyen mértékű összefonódását hangsúlyozzuk, egy humanista elmélet éppúgy kezelni képes ezeket a kérdéseket. A morális ágencia és a morális közösség további kérdéseket vet fel, különösen az intencionalitás kapcsán, mivel szerinte Verbeek nem bizonyítja meggyőzően azt, hogy a technológia aktívan hozzájárulna az emberi döntéshozatalhoz. Az erkölcsileg releváns döntésekben továbbra sem vitatható az emberi ágencia mivel, ha csak a befolyás és a hatás

¹⁷ „*Észleléseink és tapasztalataink, cselekedeteink és életmódunk, az emberi lét mindezen elemei a technológiával való szoros kölcsönhatásban alakulnak ki.*” (saját fordítás)

elendő volna, hogy valami részévé váljon a morális közösségnek, akkor ezt a jogot nem vitathatnánk természeti elemek vagy tárgyak esetében sem (Sand, 2018, pp. 33-36). Sand a saját innovációetikai kísérletében az emberi ágens jelentőségét hangsúlyozza és egy erényetikai megközelítést alkalmaz:

„The ethics of innovation developed here analyzes the basic conditions for responsible agency, taking into account the (often limited) range of possible alternatives available to agents, the ways in which they cultivate dispositions and excellences while navigating the maze of innovations to find morally desirable pathways and discover solutions (which constitutes another excellence of intelligence: creativity), and how they harmonize their personal passions and interests with the external demands of morality”¹⁸ (Sand, 2018, p. 41).

Egy az innovációval kapcsolatos normatív elméletnek olyan etikai keretrendszert kell biztosítania – írja, amely iránymutatást ad az új technológiák felelősségteljes fejlesztéséhez. Ennek a keretrendszernek meg kell határoznia azokat az erényeket és jellemvonásokat, amelyekre az innováció folyamatában részt vevő szereplőknek – az innovátoroknak: fejlesztőknek, kutatóknak - törekedniük kell annak érdekében, hogy a technológiai újítások megfeleljenek az uralkodó etikai alapelveknek (Sand, 2018, p. 5). Az innovátor személyes képességei, elköteleződése bizonyos szempontból helyettesíti a szervezeti szintű normatív elköteleződést, mivel az egyén dimenziójába helyezi át a felelőség súlyát, az ő cselekvési képességei, reflexiója és diszpozíciói határozzák meg az innovációs folyamat etikai kereteit.

A felelőség kérdését Vincent Blok az általa *episztemikus elégtelenségnek* nevezett jelenségben látja még tovább bonyolódni: egy-egy problémakörrel (mint például a globális felmelegedés) birtokolt tudásunk lényegében nem elégséges ahhoz, hogy képesek legyünk eldönteni, hogy egy innováció végeredményét tekintve milyen besorolást kapjon, jó vagy potenciálisan káros fejlesztésként kategorizáljuk. A felelőség ennek a tükrében értékelődik újra: felelőségre vonhatunk-e bárkit a fejlesztése jövőbeli hatásaiért úgy, hogy nincs tényleges tudása azok bekövetkeztének valószínűségéről (Blok, 2022, p. 901)? A felelőség azonban nem csak az innovációs folyamat szereplőire, mint személyekre korlátozódik: a kutatólaboratóriumok, vállalatok és általánosságban az innováció kereteit biztosító szervezetek morális ágenciája is fontos. Ez azonban a vállalkozói ethosz sajátosságát jelentő kockázatvállalás problematikájához vezet minket, mivel a piaci értelemben vett

¹⁸ „Az itt bemutatott etika az erkölcsileg felelős cselekvés alapvető feltételeit elemzi, figyelembe véve a cselekvők számára elérhető (gyakran korlátozott) lehetséges alternatívák körét, azokat a módokat, ahogyan jellemvonásaikat és erősségeiket fejlesztik, miközben az innovációk útvesztőjében navigálnak, hogy erkölcsileg helyes utakat találjanak és megoldásokat fedezzenek fel (amely az értelem egy újabb erőssége: az alkotókészség), és azt, hogyan hozzák összhangba személyes szenvedélyeiket és érdekeiket az erkölcs külső követelményeivel.” (saját fordítás)

kockázatvállalás megítélése az üzleti etika szempontjából egy elfogadható és szükséges elv, amennyiben figyelembe veszi az érintettek kitettségét és beépít kockázatkezelési mechanizmusokat.

Blok elmozdul az innováció fejlesztésként, kvázi találmányként való értelmezésétől és annak az általam is preferált gyakorlati, *multidimenzionális folyamatként* történő felfogását hangsúlyozza, amelyben az innováció „*új vagy jelentősen továbbfejlesztett elméletek, termékek, szolgáltatások, technológiák, folyamatok, szervezeti formák vagy piaci megközelítések létrehozása, adaptációja, implementációja és diffúziója*” (saját definíció - lásd a dolgozat második fejezetét).

Az innováció cselekvésalapú megközelítése azért jelent váltást etikai szempontból, mert az innovációnak a megismerésre, kognitív tevékenységekre redukálódó szemléletéből kilépve, egy akcióelmélet alapján alakítja ki az innovációetikáját, ahol a tudás és a szándékok másodlagosak és a fókusz a viselkedésre és a tevékenységre irányul. Blok kidolgoz egy olyan innovációetikai keretet, amely négy jellemző mentén határozza meg az innovációs folyamatok irányítójának *ethosát*. Először is az egyetemes elveknek és normáknak való megfelelés mellett az innovációs gyakorlatoknak reagálniuk kell a társadalmi környezet felől érkező konkrét igényekre, az etikus cselekvés ezen igényekre és szükségletekre adott válasz: helyzetfüggő és az aktuális problémákra irányuló cselekvés. Az innovációs folyamatban részt vevő szervezeteknek belső kritikai folyamatokat is be kell építeniük, gyakorlataiknak nem a külső szereplők meggyőzésére kell koncentrálniuk, hanem nyitottnak kell maradniuk a kívülről érkező kritikára. Végül negyedikként egy *sebezhetőségi stratégiát* (vulnerability strategy) javasol a társadalmi részvétel erősítésére, mert a saját episztemikus hiányosságainkból fakadó elfogultságunk megnehezíti az innovációk jövőbeli hatásaink előrejelzését, ennek érdekében erősíteni kell a társadalmi részvételt és reflexiót (Blok, 2022, pp. 906-908).

Michel Bourban és Johan Rochel az innovációetikát (IE) a felelős innováció (RI) felől közelíti meg, két fontos kérdést kiemelve. Az egyik a fogalmi határvonalak újrarajzolása, vagyis figyelembe kell vennünk, hogy az innováció megváltoztatja az alapvető kategóriákat, átértelmezve az általunk használt fogalmakat (pl. természetes-mesterséges) és ennek normatív vonatkozásai is vannak. A másik az új felelősségi formák bevezetése, mivel meglátásuk szerint a technológiai innováció átalakítja a felelősség klasszikus értelmezését: az új fejlesztések kollektív tevékenységek eredményei, de mögöttük egyéni szándékok, motivációk állnak és összetettségük révén számos eddig nem vizsgált szempontot szükséges beemlíteni az értékelésükbe (Bourban & Rochel, 2019, pp. 386-388). Az általuk bevezetett normatív

keretrendszer szerint a felelősség számos szereplő tevékenységének eredménye, összegző jellegű (aggregatív) megközelítés, kiterjed a nem humán létformákra is (inkluzív) és mély bizonytalanság jellemzi – ezt teszi teljeskörűvé (total responsibility). Ezt a szélesebb felelősségfogalmat javasolják a felelős innovációval kapcsolatos diskurzusban, a következményelvű megközelítés mellett az erényetikai szempontok figyelembevételével. A szerzők célja, hogy egy önálló innovációetikát alkossanak a gyakorlati és autonóm alkalmazott ágazati etikák

A gyakorlati etikai döntéshozatalt támogató általános konszenzusereső kísérletekre kiváló példa Van De Poel és Royakkers etikai ciklus (The Ethical Cycle) modellje: *„A tool in structuring and improving moral decisions by making a systematic and thorough analysis of the moral problem, which helps to come to a moral judgment and to justify the final decision in moral terms”*¹⁹ (Van De Poel & Royakkers, 2011, p. 137). A szerzők 2007-es tanulmányukban rávilágítanak arra, hogy az alkalmazott etika, mint morális elvek vagy elméletek alkalmazása konkrét szituációkra számos problémát vet fel, mert nincs egyetlen általánosan elfogadott etikai elmélet vagy keretrendszer, amely minden esetben releváns lenne. Az etikai fogalmak bonyolultak, eltérő morális nézetrendszereket tükrözhet az értelmezésük és a használatuk. Az etikai elméletek elsődlegesen heurisztikus természetűek, segítenek a morális kérdések megfogalmazásában és a döntéshozatalban, a fő feladatuk éppen ebben rejlik: érveket biztosítanak a végső ítéletekhez (Van De Poel & Royakkers, 2007, pp. 2-3). Emiatt érdemes egy általános keretrendszer elfogadása helyett többféle etikai nézőpontot, elméletet és alapelveket figyelembe venni a mérlegelésnél. Az etikai ciklus modell ez alapján épül fel: [1] első lépésben megtörténik a morális probléma világos meghatározása, ami magában foglalja az ágens megjelölését is; [2] a problémaelemzési szakaszban megvizsgáljuk az érintetteket és az ő érdekeiket, továbbá a rendelkezésünkre álló információkat, tényeket; a harmadik szakasz [3] a cselekvési lehetőségek azonosításáról és a stratégia kidolgozásáról szól; a negyedik szakaszban [4] történik meg az értékelés, amely különféle etikai elméletek (konzekvencializmus, deontológia, erényetika, utilitarizmus) figyelembevételével segíti az adott eset összes dimenziójának komplex feltárását. Az ötödik [5] szakaszban történik meg a reflexió, amelyben az egyén intuitív ítélete is jelentőséget kap, az etikai érvek számbavétele mellett. Ebben a szakaszban kap szerepet az argumentáció, ami a végső döntés indoklásában hangsúlyos érvek melletti állásfoglalást, illetve az nem elfogadott etikai nézetek elutasítását alapozza meg. Ennek

¹⁹ „Az erkölcsi döntések strukturálásának és javításának eszköze az erkölcsi probléma szisztematikus és alapos elemzésével, amely segít erkölcsi ítéletet hozni és a végső döntést erkölcsi szempontból indokolni.” (saját fordítás)

az ötlépcsős modellnek a végén születik meg az erkölcsi ítélet (Van De Poel & Royakkers, 2011. pp. 138-157). A modell alkalmazása során törekedni kell egy széles reflektív egyensúlyon alapuló megközelítés érvényesítésére, mivel így leszünk képesek a legváltozatosabb érvek és indokok figyelembevételére.

T.T. Lennerfors és Kiyoshi Murata a 2023-ban megjelent *Innovation Ethics* címet viselő tanulmányában a (főként technológiafókuszú) innovációetikáról szóló tudományos diskurzust elemzi, bemutatva a kortárs megközelítéseket és modelleket. A tanulmányban megfogalmazzák saját innovációetikai alapvetésüket, amelynek fő célja, hogy ösztönözzék a témáról szóló vitát, új nézőpontokkal gyarapítva azt. Fontosnak tartják, hogy az alapvetően kritikátlanul pozitív képet, amely az innovációval kapcsolatban uralkodik az egyetemeken mérnöki, üzleti és természettudományos képzéseiben tanuló hallgatók körében, egy kritikus etikai megközelítéssel árnyalják, illetve a nem technológiai képzések hallgatói körében is párbeszédet kezdeményezzenek a technológiai haladással kapcsolatosan. A modern innovációetikának szerintük öt fő kérdésre kell fókuszálnia:

- *Ösztönöznie kell a kritikus gondolkodást és el kell kerülnie az értékek tárgyiasítását*
Az innovációetika nem szabad, hogy felszínessé váljon, értékek és elvek listászerű felsorolásává – reflexív módon kell viszonyulnia az univerzális morális alapokhoz, folyamatosan párbeszédben a kulturális és társadalmi közeggel, amiben működik
- *Az újítás mellett, a meglévő értékek megőrzésére és gondozására is figyelmet kell fordítson*
Mérlegelnie kell, hogy egy adott innováció milyen hatást gyakorolhat a már létező struktúrákra és értékekre. Törekednie kell egy olyan egyensúly kiépítésére, amely a fejlődést úgy segíti elő, hogy közben nem rombolja le a társadalom vagy kultúra alapjait
- *Az innovációetika folyamatjellegű és iteratív, tematikusan nem szűkülhet le új találmányok etikai vonatkozásaira*
Az innováció fogalmának komplexitásából következően, az innovációetikának egy széles spektrumot kell felügyelnie. A fejlesztés fázisán túl figyelmet kell fordítania az adott technológia alkalmazása során felmerülő etikai kérdésekre, nyomon követve az innováció életciklusát, folyamatosan korrigálva az etikai kereteket, rugalmasan alkalmazkodva a változó körülményekhez és az új felismerésekhez
- *Nyitottnak kell lennie az innováció eltérő közegeire és kontextusaira*
Az innovációs etikának figyelmet kell fordítania a különböző közegekből érkező innovációkra, fókusza nem korlátozódhat csak az orvosi vagy technológiai szektorra.

Kritikusan kell vizsgálnia az innovációt befolyásoló erőket és érdekeket, és a társadalom domináns szektorain kívül eső területekről érkező innovációkat is tárgyalnia

- *Az etikai kérdések összetettségének elismerése mellett, gyakorlat és cselekvésorientált marad*
Az innovációs etikának egyszerre kell figyelembe vennie az etikai kérdések komplexitását és gyakorlatiasnak maradnia. A cél nem a tökéletes megoldások megtalálása, hanem, hogy törekedjen megoldások találására, a leküzdhetetlennek tűnő kihívások ellenére is (Lennerfors & Murata, 2023, pp. 48-50)

Ha a felelősség felől vizsgáljuk az áttekintett elméleteket, több kritikai észrevételt is megfogalmazhatunk. Az egyik ilyen az ember és a technológia viszonyában a *felelősség* és az *ágencia* kérdése, ami a Verbeek és Sand nézetei közötti szembenállásban mutatkozik meg. Verbeek, aki a technológiát aktív szereplőnek látja az emberi létezés alakításában, egyben túl nagy ágenciát is tulajdonít neki, megosztva a felelősséget az ember és a technológia között. Ennek kapcsán Sand joggal veti fel, hogy a technológia nem mutat szándékosságot vagy erkölcsi ítélőképességet, így a morális közösség részévé sem válhat, az etikai döntések meghozatalában az emberi ágencia megkérdőjelezhetetlen marad, a felelősség nem osztható meg nem emberi entitásokkal. Blok elméletében az *episztemikus elégtelenség* koncepciója, az innovációk jövőbeli hatásainak felmérhetetlenségére világít rá, ami felveti azt a kérdést, hogy milyen mértékben tehetők felelőssé az innovációs tevékenységben részt vevők a potenciális negatív következményekért, amennyiben nem rendelkeznek elegendő tudással ennek előrejelzésére? *A sebezhetőségi stratégia* alkalmazása lényegében megfeleltethető lenne a felelősségetika elővigyázatossági elvének, amennyiben a kockázatok felmérésének és a károk kiküszöbölésének elve normatív elvárásként szerepelne, amely esetben az elszámoltathatóság biztosított volna a fejlesztők kapcsán. Bourban és Rochel elmélete szintén az emberi ágencia és elszámoltathatóság relativizálását hozza, az egyéni felelősség jelentőségének a csökkentésével és általában a felelősség aggregatív megosztásával. A Van De Poel és Royakkers által javasolt modell, habár strukturált és reflektív megközelítést kínál az erkölcsi döntéshozatalhoz, de nem kezeli rugalmasan az innováció által felvetett egyedi kihívásokat. Lennerfors és Murata sem ad iránymutatást arra vonatkozóan, hogy hogyan rangsoroljuk az egymással versengő értékeket, vagy hogyan oldjuk meg a gyakorlatban felmerülő etikai dilemmákat.

Figyelembe véve az itt vizsgált megközelítéseket, levonható az a következtetés, hogy az innovációetika mint ágazati etika, nem rendelkezik más alkalmazott területekhez (az orvosi vagy a mérnöki etikához) hasonló megszilárdult elméleti keretekkel, sok esetben egymással összemérhetetlen elvek és elméletek versengése jellemzi (például az erős technológiai fókusz

eltereli a figyelmet az innováció más dimenzióiról). Az eltérő filozófiai hagyományok (erényetika, utilitarizmus) elemei éppúgy megjelennek egyes modellekben, mint metaetikai diskurzusok, ami a gyakorlati alkalmazhatóságot gyengíti. Az erkölcsileg felelős cselekvés normatív alapjainak kijelölésében megjelent a törekvés a konszenzusra (ezt részletesen látjuk majd az ötödik fejezet esettanulmányában), minden szerző próbál irányelveket alkotni arra vonatkozóan, hogy az innovációs folyamat során mi az erkölcsileg helyes cselekvés, azonban az elméletek adósak maradnak az alkalmazás módszereivel. Azekere a hiányosságokra úgy gondolom, hogy egy felelősségetikai szemlélet hangsúlyosabb alkalmazása jelentene megoldást – ahogy emellett a következőkben érvelni fogok.

4.3 A felelős kutatás és innováció etikai keretei: közösségi és egyéni perspektívák

Természetesen nehéz egy átfogó, adaptív etikai keretrendszer nélkül gyakorlati útmutatókat gyártani, ez a probléma, amivel a felelős innováció (Responsible Innovation – RI) is szembesülni kénytelen. A felelősségteljes kutatás és innováció (Responsible Research and Innovation, RRI) egyfajta hidat képez az elméleti megfontolások és a gyakorlati megvalósítás között. Egy olyan megközelítést takar, amely lehetővé teszi az érintettek számára, hogy már a fejlesztési folyamatok korai szakaszában részt vegyenek releváns ismereteket szerezve, értékelhessék a fejlesztések potenciális következményeit és a rendelkezésre álló lehetőségeket a társadalmi szükségletek és erkölcsi értékek szempontjából, és ezeket a szempontokat beépítsék a kutatás-fejlesztési folyamat tényleges követelményei közé (European Commission, Directorate-General for Research and Innovation, 2013, p. 3).

De mitől lesz az innováció felelős? Mielőtt ezt megvizsgálánk, érdemes áttekintenünk, hogy milyen esetekben beszélhetünk felelőtlen (irresponsible) innovációról – ebben Von Schomberg segít minket – aki négy kategóriát sorol fel, kiemelve, hogy az esetek többségében ezek a dimenziók vegyesen vannak jelen. A *technológiai nyomás* jellemzően valamilyen vállalati szereplő gazdasági érdekből fakadó befolyásgyakorlása az innovációs folyamatra; az *alapvető etikai elvek elhanyagolása* a fejlesztési folyamat során figyelmen kívül hagyott morális kérdéseket és potenciális problémákat jelenti; a *szakpolitika húzóereje* amikor a szakpolitikai döntések válnak az innováció fő mozgatórugóivá, a fejlesztők próbálnak alkalmazkodni a szabályozási környezethez vagy kihasználni a kormányzati támogatások által kínált lehetőségeket és végül az *elővigyázatossági intézkedések és a technológiai előrelépések hiánya* azokat a fejlesztési folyamatokat jelenti, amely elővigyázatossági elvek figyelembevétele nélkül, elégtelen hatásvizsgálati eredmények birtokában történik meg a fejlesztés (Von Schomberg, 2013, pp. 60-63).

A felelős innováció ezekkel szemben egy kezdeményező hozzáállást kíván meg a technológiai átalakulással és fejlődéssel járó kockázatok és bizonytalanságok kezelésére, miközben megteremti a feltételeit annak, hogy az innováció előnyeiből minden társadalmi réteg igazságosan részesüljön (Stilgoe et al., 2013). Ezt emeli ki az European Commission: Directorate-General for Research and Innovation szakpolitikai összefoglalójának definíciója: „Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products”²⁰ (Von Schomberg, 2011, p. 9). A dokumentumot szerkesztő Von Schomberg szerint az innovációs folyamatba való korai társadalmi bevonódás segíti a fejlesztések társadalmi elfogadását, illetve a későbbi hatások szabályozására is pozitív hatással lehet. Elválasztja egymástól az innováció *termék* és a *folyamat* dimenzióját – az utóbbit egy multidiszciplináris szemléletben minden érintett és érdekelt közösen alakítja: az innovátorok érzékenyebbé válnak a társadalmi szükségletekre, míg a társadalmi szereplők konstruktív kritikáikkal és észrevételeikkel segítik egy közös felelősségvállalás kialakítását. A lenti táblázat ezeket foglalja össze:

1. táblázat. A felelős kutatás és innováció jellemzőinek áttekintése (Von Schomberg, 2011)

A FELELŐS KUTATÁS ÉS INNOVÁCIÓ JELLEMZŐI	
TERMÉK DIMENZIÓ	FOLYAMAT DIMENZIÓ
Az elővigyázatossági intézkedések és a technológiai előrejelzések intézményesítése	Magatartási kódexek használata
Az elővigyázatosság elvének alkalmazása; folyamatos kockázatértékelés; folyamatos nyomon követés	Az elszámoltathatóság biztosítása: szabványok, tanúsítási rendszerek, címkék használata
Demonstrációs projektek alkalmazása: a kockázat azonosításától az innováció irányításáig	Az etika, mint a technológia tervezési alapelve
	Normatív irányítási modellek
	Folyamatos nyilvános párbeszéd: a szakpolitikai húzóerő és technológiai nyomás kezelése

A felelős innováció jelentősége az ellenőrizetlen technológiai fejlődésből adódó nem szándékolt következmények és potenciális károk kezelésében rejlik – a történelem során számos példával találkozunk, olyan kezdetben pozitív áttörésként üdvözölt újításokkal,

²⁰ „A felelős kutatás és innováció olyan átlátható, interaktív folyamat, amelynek során a társadalmi szereplők és az innovátorok kölcsönösen odafigyelnek egymásra, tekintettel az innovációs folyamat és a piacra kerülő termékek (etikai) elfogadhatóságára, fenntarthatóságára és a vele kapcsolatos társadalmi igényre.” (saját fordítás)

amelyek a későbbiekben mégis árnyoldalaik miatt híresültek el. Edward Tenner történész (1996) számos példát hoz olyan találmányokra, megoldásokra, technológiákra, amelyek nem szándékolt és gyakran negatív következményekkel jártak, a növényvédő szerek környezeti hatásaitól kezdve az automatizálás okozta társadalmi és gazdasági károkig.

Owen és munkatársai négy dimenzióban határozzák meg, hogy milyen a felelős innováció: [1] *előrelátó*, azaz számol a potenciálisan nem tervezhető hatások következményeivel; [2] *reflektív*, figyelmet fordít és reflektál a dilemmákra, kockázatokra és mögöttes célokra; [3] *deliberatív* vagyis párbeszédkész, nyitott a különböző érdekeltek perspektíváinak megvitatására és beépítésére; továbbá [4] *reszponzív*, vagyis a felelős innováció egy folyamatos, rugalmas és részvételen alapuló adaptív tanulási folyamat, amely képes alakítani és irányítani az innováció irányát (Owen et al., 2013, p. 38).

Nem vitatható, hogy a tudományba és technológiába vetett közbizalom fenntartásában is fontos szerepet szánunk tudományos és döntéshozó fórumok a felelős innovációnak. Az egyre összetettebbé váló innovációs tevékenységek következményei egyre szélesebb körűek, ennek függvényében a közvélemény figyelve és ellenállásának kockázata is nő (Gaskell et al., 2004). Az eltérő igények, nézőpontok és etikai elvek iránti elkötelezettségek figyelembevételével a felelős innováció eredményesen szolgálhatja a közbizalom kiépítését (von Schomberg, 2013).

Annak ellenére, hogy nyomós érvek szólnak a felelős innováció szükségessége mellett, – kiemelve, hogy a kockázatok kezelésének egy lehetséges modelljét jelenti csupán, amely egy fenntarthatóbb és társadalmilag igazságosabb irányba tudja elmozdítani az innovációs tevékenységeket (Guston, 2014) – összességében mégis azt kell látnunk, hogy nem jelent megoldást a technológiai fejlődés valós lételméleti-episztemológiai problémáinak értelmezésében és kezelésében. Egyesek szerint a technológiai rendszerek inherens bizonytalansága és komplexitása lehetetlenné teszi a valószínűsíthető következmények széles spektrumának kezelését (Jasanoff, 2016), így a felelős innovációra tett minden kísérlet végső soron hiábavaló, mivel a nem szándékolt következmények mindig meg fogják haladni az azok kezelésére álló lehetőségeinket és képességeinket. Mint látni fogjuk, erre kínál stratégiát a Hans Jonas által bevezetett félelem heurisztikája, amely egyfajta jelzőrendszerként segíti a kutatókat a munkájuk során, hogy az egyes fejlesztések negatív hatásaitól való félelmet is kalkulálva, a munkájuk korai szakaszában azonosítsák a potenciális kockázatokat (Jonas, 1984). Vannak szerzők, akik szerint, maga a technológiai fejlődésre való törekezés elhibázott és lényegében fenntarthatatlan. Eugene Thacker például úgy érvel, hogy mivel az ember nem ismeri fel a saját tudásának a korlátait és a komplex rendszerek kiszámíthatatlan voltát, a törekvése, hogy a

világot uralja végül önpusztításhoz vezet. A felelős innováció így csupán egy önáltatás, és hangzatos, de üres közhely, mert a technológiai fejlődés által felvetett lényeges filozófiai kérdésekkel nem tud és nem is képes foglalkozni azok ontológiai mélységében (Thacker, 2011).

A disszertációban amellet érvelek, hogy a felelős innováció akkor is nélkülözhetetlen eszköz a nem szándékolt következmények mérséklésére és a kockázatok kezelésére, ha a fejlődés által felvetett filozófiai dilemmákra adandó válaszokkal adós is marad. Azon etikai megfontolások, amelyek alátámasztják a felelős innovációt, a technológiai rendszerek eredendő bizonytalanságának felismerésén alapulnak, így a modell határait is kijelölik. A bizonytalanság áthidalását egy holisztikus szemlélet segítheti, a tudásunk korlátainak elismerése mellett, továbbra is hangsúlyt helyezve a tanulás és az adaptáció folyamataira: ehhez a felelősségetika kínál keretet.

4.4 Az etika, az innováció és a felelősség összefüggései a kortárs tudományosságban

A filozófiai elméletek relevanciája és alkalmazhatósága nagyban függ attól, hogy milyen gyakorisággal, milyen széles körben és milyen kontextusban jelennek meg a tudományos diskurzusban. Az elméleti áttekintést követően szeretném megvizsgálni, hogy a dolgozat tárgyát képező témakörök – felelősség, innováció, etika – mennyire vannak jelen a kortárs tudományos érdeklődésben, kiemelt figyelemmel a tudományterületi fókuszokra. Szükségszerű egy statisztikai adatokkal alátámasztott fejezet beiktatása, hiszen ezekkel a számszerűsített elemzésekkel lehet demonstrálni, hogy egyes fogalmakkal milyen kontextusokban találkozunk a leggyakrabban (hogyan kapcsolódnak össze egyes tudományterületekkel), milyen követhető változások vannak a publikációk számában (ami a téma iránti érdeklődés jelzőszáma) és egyes fogalmi kapcsolódások egymással milyen viszonyban írhatók le. A tudományos publikációk trendjeiről és fókuszairól alkotott kép egy-egy kutatási terület jelentőségéről is árulkodik, bemutatva, hogy bizonyos kérdések nem csupán elméleti spekulációk tárgyai, hanem aktív kutatási területek. A fejezetben vizsgálom, hogy a dolgozat tárgyát képező etika, innováció és felelősség kontextusa mennyire kap figyelmet, illetve milyen mértékben reflektál a tudományos közösség az új technológiai és társadalmi kihívásokra – ami az esettanulmányom tárgyául választott mesterséges intelligencia relevanciáját hivatott indokolni.

Az alkalmazott etikai kutatásokban az empirikus adatok elemzése, felhasználása és általánosságban az empiria integrálása sok szempontból gazdagítja és pontosítja az elméleti vizsgálatokat, jelentősen növelve az etikai elemzések relevanciáját, hatékonyságát és adaptálhatóságát. Az empirikus megközelítés révén történik meg az etikai kérdések

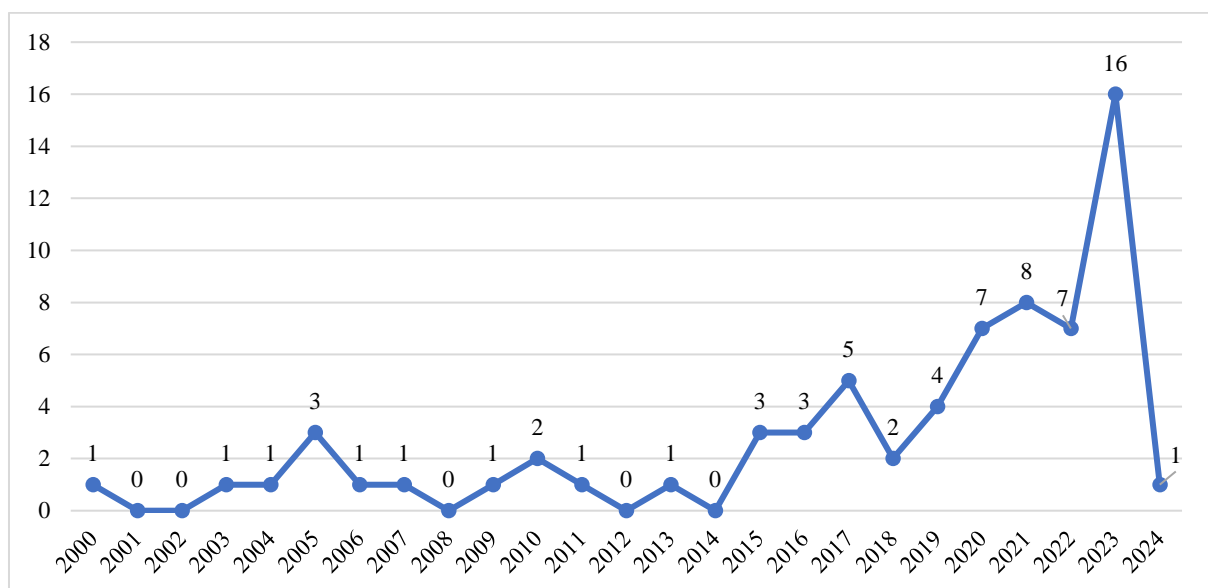
kontextualizálása, amelyek révén az egyes területek komplexebb megismerésére és megértésére nyílik lehetőség. Mindezek mellett az empirikus kutatás lehetőséget nyújt az etikai elméletek pontosítására és validálására, ami egy iteratív folyamatban segíti elő területspecifikusan alkalmazható etikai modellek kidolgozását. Az elmúlt évtizedekben, különösen bioetikai kérdésekben nagy érdeklődés övezte a különböző tudományterületek – jellemzően az élettudományok, a filozófia és a társadalomtudományok – együttműködését (Sugarman, 2004; Weisz, 1990). Az empirikus kutatási módszerek alkalmazása több területen elterjedté vált, az etika hagyományosan elméleti kereteiből kilépve egy valóban hibrid tudományosság gyakorlati megvalósítása felé. Az alkalmazott etika ezen empirikus orientációja, a „világot úgy igyekszik leírni ahogy van, míg a teoretikus megközelítés úgy, ahogy kellene lennie” –, a két megközelítés közötti feszültség feloldásának a megoldása az empirikus adatok felhasználásának elfogadása a megfelelő normatív válaszok kidolgozásához (Leget et al., 2009, s.p.).

Bibliometriai elemzési megközelítést alkalmazva vizsgálom az innovációs etikával kapcsolatos publikációk megjelenését, továbbá ezek trendjeit és tendenciáit. A bibliometria – mint a tudományos publikációk statisztikai elemzésére szolgáló módszertani keret – lehetővé teszi, hogy számszerűsíthető adatokat nyerjünk tudományos szakirodalmi adatbázisokból, ezeket elemezzük és vizualizáljuk (van Eck & Waltman, 2014). A tudományos világ több nagy adatbázist is létrehozott a hivatkozások nyomon követésére és elemzésére, ezek gyűjtik és rendszerezik a hivatkozási adatokat, amelyekből aztán különböző mérőszámokat állítanak elő. A bibliometriai elemzéshez adatforrásként a Scopust választottam, mint az egyik legnagyobb absztrakt- és idézettségi adatbázist, amely több mint 7000 kiadóval áll kapcsolatban, 27950 peer-reviewed szakfolyóiratot indexel és 98%-os pontosságú adathozzáférést biztosít. A bibliometriai elemzések elvégzéséhez a Scopus keresés funkcióját használtam, ami lehetővé tette a keresési eredmények komplex vizsgálatát olyan kritériumok alapján, mint a dokumentumok évenkénti eloszlása, ezen túlmenően vizsgáltam még a szerzőket, az intézményeket, és a dokumentumtípusokat is. A vizsgálati stratégiám úgy alakítottam ki, hogy az innováció etikai dimenzióival kapcsolatos publikációk legszélesebb körét meg tudjam találni, ugyanakkor a téma specifikumát is megőrizzem.

Első lépésben a kutatási témámhoz kapcsolódó kulcsszavak kiválasztásával kezdtem, így a keresési kifejezések közé került az „etika”, az „innováció” és a „felelősség”. A tág keresések kezelhetetlen számú találatot eredményeztek, emiatt módosítanom kellett többször a kifejezéseket és úgynevezett a logikai operátorokat – például az AND és OR kötőszavakat – használtam a keresési kifejezések hatékony összekapcsolására. A keresést leszűkítettem

tanulmányokra (article), konferenciaközleményekre (conference paper), könyvfejezetekre (book chapter) és könyvekre (book). Mivel a Scopus adatbázis 1996-tól kezdte szisztematikusan indexelni az adatbázisába bekerülő folyóiratokat, illetve bővítette az indexelt tudományágak körét is, ezért vizsgálatomban az 1996 és 2024 közötti időszakban megjelent publikációkat vizsgáltam csak²¹. A Scopus „Analyze Search Results” (Keresési eredmények elemzése) funkcióját használtam a visszakeresett adatok vizsgálatához, ez az eszköz lehetővé teszi a keresési eredmények mélyreható elemzését számos kritérium alapján (I. melléklet 1. táblázat). Meg kell jegyezni, hogy a vizsgálat időpontjáig érvényesek a 2024-es adatok, azaz jelen esetben ez 2024. június 30-ig.

A vizsgálat első körében az *innovációetika* és az *innováció etikája* ("innovation ethics" OR "ethics of innovation") kulcsszavak előfordulását vizsgáltam, és alig 70 dokumentumot találtam a 2000-2024 közötti időszakban (a korábbi évekre nem adott a keresés találatot). A találatok csekély száma a terület relatív újdonságának tudható be, az innovációetika mint ágazati etika még nem artikulálódott, önálló tudományos folyóirattal nem rendelkezik. Az alacsony elemszámtól eltekintve, jól követhető a témakör iránt exponenciálisan növekvő figyelem (2. ábra).

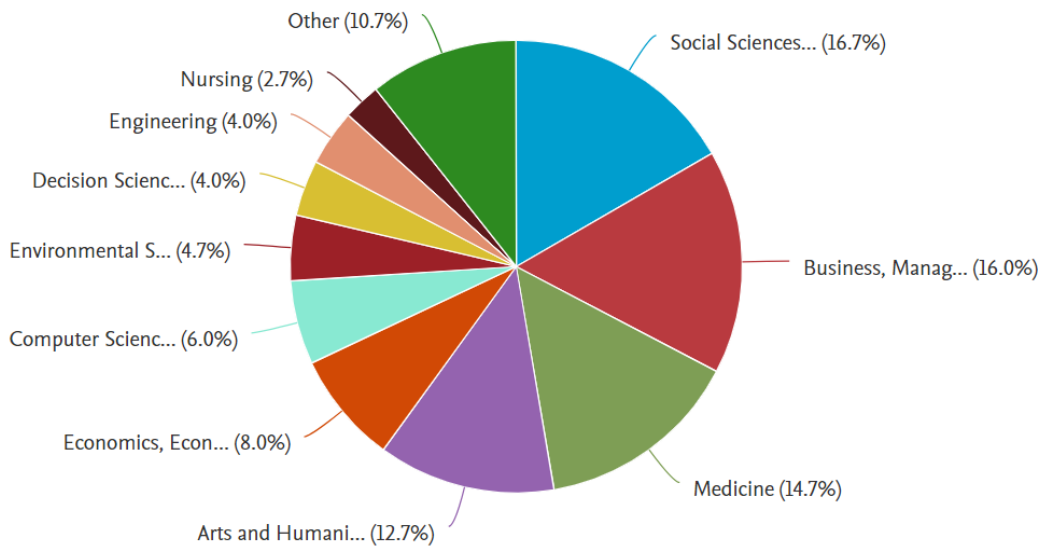


2. ábra. Az *innovációetika* és az *innováció etikája* kulcsszavak előfordulása (db) (forrás: Scopus)

Az, hogy a megjelent tanulmányok miként oszlanak meg az egyes tudományterületek között, egy multidiszciplináris érdeklődésre utal, továbbá az innováció több tudományterületre

²¹ A dolgozat I. melléklete tartalmazza a részletes adatokat, a keresési karakterláncot, illetve további táblázatokat és grafikonokat. A szövegkohézió megőrzése okán csak néhány, jelen fejezet szempontjából releváns dimenziót szeretnék kiemelni a törzsszövegben.

kiterjedő hatókörére. Kiemelkedő érdeklődést láthatunk a társadalomtudományok (16,7%) illetve az üzlet és menedzsmenttudomány területén (16,0%), ami az innováció társadalmi és gazdasági jelentőségét támasztja alá, továbbá az orvostudományokban (14,7 %) mint az etikai kérdésekre érzékeny innovációvezérelt diszciplínában (3. ábra).



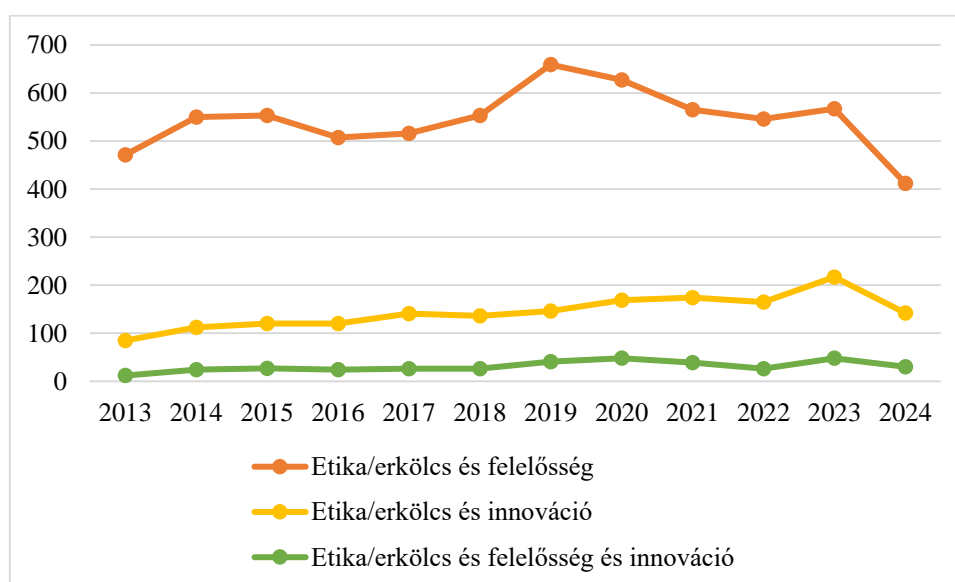
3. ábra. A megjelent tanulmányok tudományterületi megoszlása (innovációetika) (forrás: Scopus)

Ezután megvizsgáltam az etika/erkölcs és különböző kulcsszavak kapcsolódását, úgy, mint az 'etika/erkölcs és felelősség', az 'etika/erkölcs és innováció', valamint az 'etika/erkölcs és felelősség és innováció' kulcsszavak kapcsolatait is. Ezen kulcsszavak és kategóriák találatainak száma egy összesített (2. táblázat) táblázatban láthatók. Ki kell emelnem, amikor a találatok számát említem, akkor minden esetben a 2013 és 2024 közötti összesített adatszámra értem.

2. táblázat. A publikációk megoszlása évek és kategória szerinti bontásban (db) 2013-2024 (saját szerkesztés) (forrás: Scopus)

Év	Megjelent tanulmányok száma (db)		
	Etika/erkölcs és felelősség	Etika/erkölcs és innováció	Etika/erkölcs és felelősség és innováció
2024	412	142	30
2023	567	217	48
2022	546	165	26
2021	565	174	39
2020	627	169	48
2019	659	146	41
2018	553	136	26
2017	516	141	26
2016	507	120	24
2015	553	120	27
2014	550	112	24
2013	471	85	12

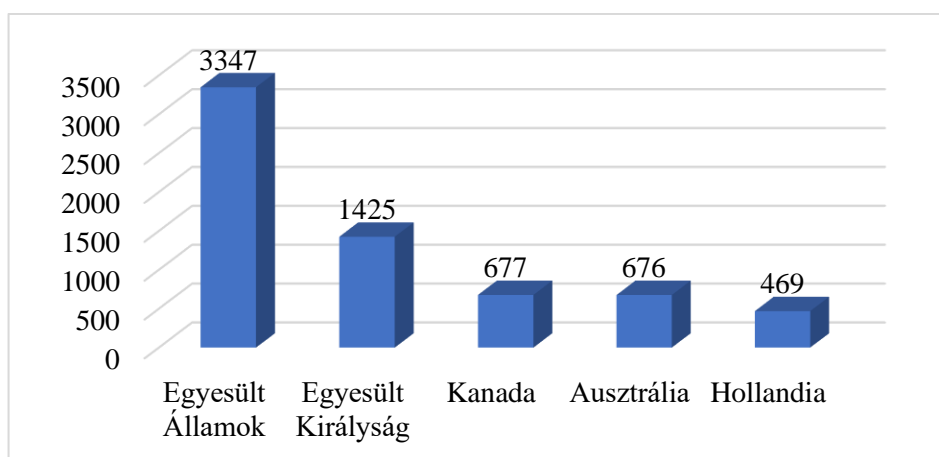
Az értékek változását egy folyamatábra is bemutatja, amelyből ténylegesen leolvasható, hogy minden esetben egyfajta növekedés figyelhető meg, azonban a kulcsszavak egymáshoz viszonyított találatai is egyértelműen láthatóvá válnak, amelyek elemzésére a későbbiekben részletesen rátérek (4.ábra).



4. ábra. A publikációk megoszlása évek és kategória szerinti bontásban (db) 2013-2024 (saját szerkesztés) (forrás: Scopus)

Az 'etika/erkölcs és felelősség' (ethics OR moral AND responsib*) kulcsszavak egy sokkal jelentősebb találati számot eredményeztek, összesen 10567 (1996-2024 között megjelent) dokumentumot tartalmaz az adatbázis. A magas elemszám mellett is egy jól követhető exponenciális növekedés látható (4. ábra), ami azt jelzi, hogy a felelősség és az etika kérdése egyre nagyobb szerepet kap a tudományos diskurzusban, és egyre több kutató foglalkozik ezekkel a kérdésekkel. Említésre méltó még, hogy kiemelkedően magas volt a COVID-időszakban (2020-22) megjelent publikációk száma.

Ha megvizsgáljuk a publikációk földrajzi megoszlását, akkor egy egyértelmű angolszász dominancia rajzolódik ki, ami rávilágít arra, hogy az ilyen jellegű tudományos kutatások döntően néhány országban összpontosulnak. Az etika/erkölcs és felelősség kulcsszavak esetén az adatokat vizsgálva látható, hogy az Egyesült Államok vezető szerepet játszik, amelyet az Egyesült Királyság, Kanada, Ausztrália és Hollandia követ (5. ábra).



5. ábra. Publikációk megoszlása földrajzi megoszlás szerint – az első 5. ország (db) (etika/erkölcs és felelősség) (forrás:Scopus)

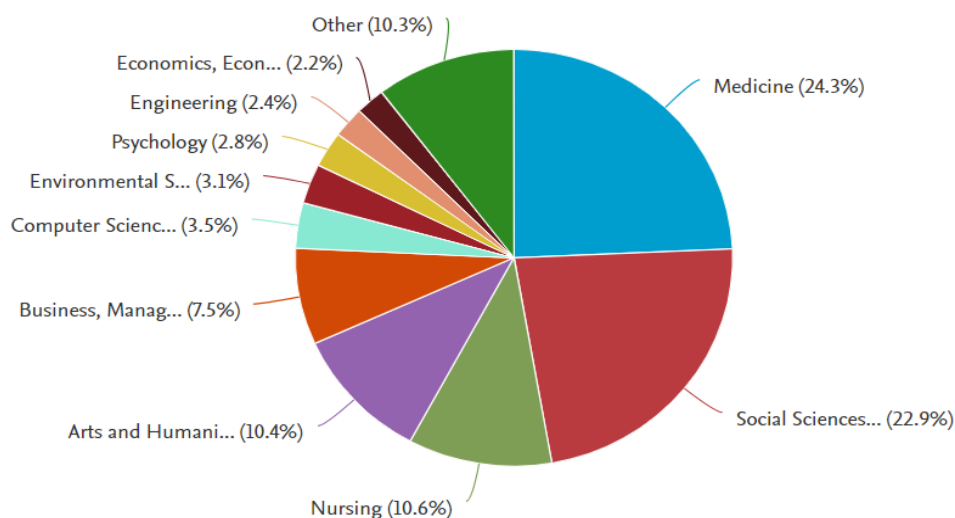
Ezek az országok nemcsak a legtöbb megjelent publikációval rendelkeznek, hanem a szakirodalmi kutatómunkám alapján a tudományos diskurzus alakításában is jelentős szerepet töltenek be. Ez részben a hagyományosan erős akadémiai háttérre vezethető vissza, továbbá az innovációban és a kutatásfinanszírozásban is ezek az országok járnak élen, az Egyesült Államok és az Egyesült Királyság vezetők a technológiai innováció terén. A továbbiakban vizsgálandó keresőszavak esetében hasonló megoszlásokat találtam, ezeket az adatokat az I. melléklet 2. táblázata mutatja be. A találatok nyelvi sajátosságait elemezve is megerősítést nyer a már említett angolszász túlsúly: az angol nyelvű publikációk száma messze meghaladja a többi nyelven megjelent publikációkét, ami a fenti földrajzi eloszlást tükrözi, másfelől alátámasztja azt a tényt, hogy az angol a tudományos világ globális lingua franca-ja (3. táblázat). A tudományos közösség nemzetközi kommunikációjának túlnyomó része angolul zajlik, és a

legnagyobb elismert tudományos folyóiratok is angol nyelvűek. Akárcsak az országok esetében, a többi kulcsszavas keresés kapcsán is ugyanezeket nyelvi arányokat találtam.

3. táblázat. A SCOPUS adatbázisban szereplő dokumentumok nyelvi megoszlása (saját szerkesztés)

A publikáció nyelve	Megjelent publikációk száma (db)
Angol	9,780
Spanyol	306
Francia	198
Német	151
Portugál	91

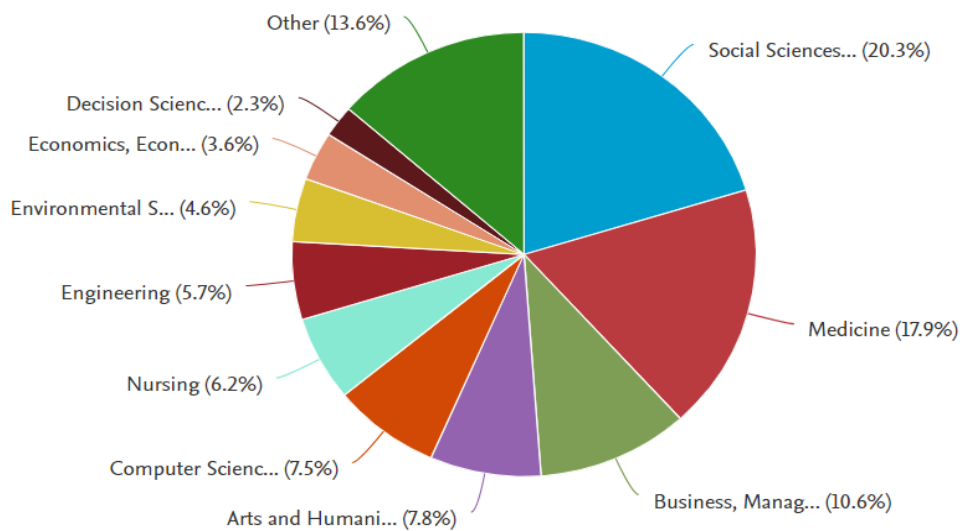
Ha az 'etika/erkölcs és a felelősség' kulcsszavakhoz kapcsolódó publikációknak a tudományterületi megoszlását vizsgáljuk, akkor azt látjuk, hogy innovációetikához képest (3. ábra) eltérést mutat (jelentősen magasabb az elemszáma), és itt egyértelműen az orvostudományoké és az élettudományoké a vezető szerep (6. ábra).



6. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és felelősség) (forrás: Scopus)

Tovább folytatva a vizsgálatot az 'etika/erkölcs és innováció' (ethics OR moral AND innovation) kulcsszavakra szűkítve, a keresés 2369 találatot hozott és az előzőhöz hasonlóan itt is egy jelentős idősoros növekedést láthatunk (2.táblázat és 4. ábra), azonban látható, hogy a találatok száma szignifikánsan alacsonyabb az etika/erkölcs és felelősség kategóriához hasonlóan. Ha a tudományterületi megoszlást vizsgáljuk láthatjuk, hogy az etika/erkölcs és felelősség kulcsszavakhoz hasonló megoszlást találhatunk, azaz az élettudományok és az

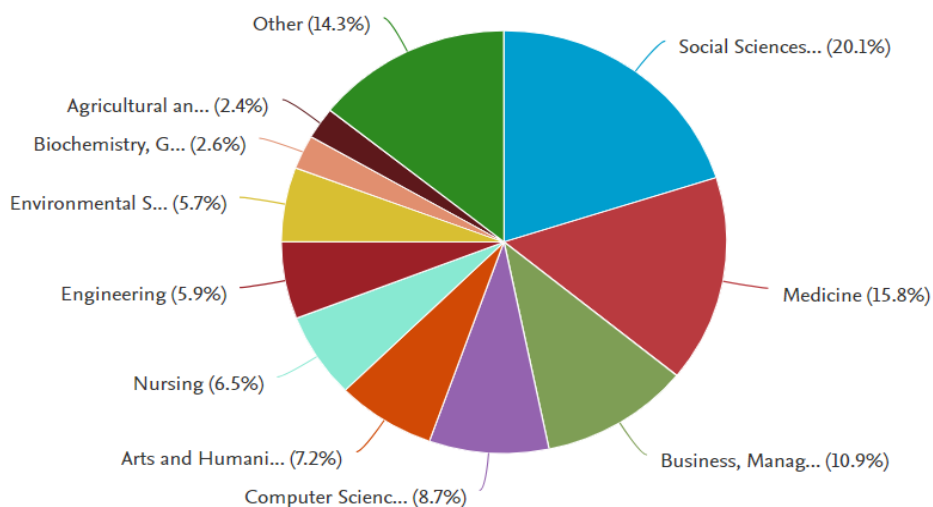
orvostudományok állnak az első kettő helyen, jelentős százalékkal. Ezt ebben az esetben a 7. ábra szemlélteti.



7. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és innováció) (forrás: Scopus)

A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és innováció) (forrás: Scopus)

Ha az 'etika/erkölcs és felelősség és innováció' (ethics OR moral AND responsib*AND innovation) kulcsszavakat vizsgáljuk együttesen, akkor elmondható, hogy összesen 458 publikációt találunk. Ehhez a kulcsszavakhoz tartozó publikációk alakulását szintén a 2. táblázat és a 4. ábra mutatja be. Azonban itt is ki kell emelni a tudományterületi megoszlást (8. ábra).



8. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és felelősség és innováció) (forrás: Scopus)

Az adatok alapján elmondható, hogy az etikai kérdések és felelősség kérdései széles körben kutatottak, különösen az orvostudományokban és élettudományokban, ahol a technológiai innovációk hatására jelentős etikai dilemmák merülhetnek fel. Az innovációhoz kapcsolódó etikai és felelősségi kérdések egyre inkább előtérbe kerülnek, de ezek összetett megközelítést igényelnek, ami a publikációk tudományterületi diverzitásában is megnyilvánul. Látható, hogy mind a három kulcsszókapcsolatnál az orvostudomány és a társadalomtudomány dominál, illetve az első hat kategóriában ugyanazok a szakterületek jelennek meg. Ezt a 4. táblázat is bemutatja. Azonban ki kell emelni az összehasonlításban, hogy a legkomplexebb és egyben legújabb kutatási irány, amely az etika, felelősség és innováció kérdéseit együtt vizsgálja, az eredmények alapján megállapítható, hogy még mindig kevésbé kutatott, mint a többi vizsgált kapcsolat, ami a publikációk alacsonyabb számában is megmutatkozik.

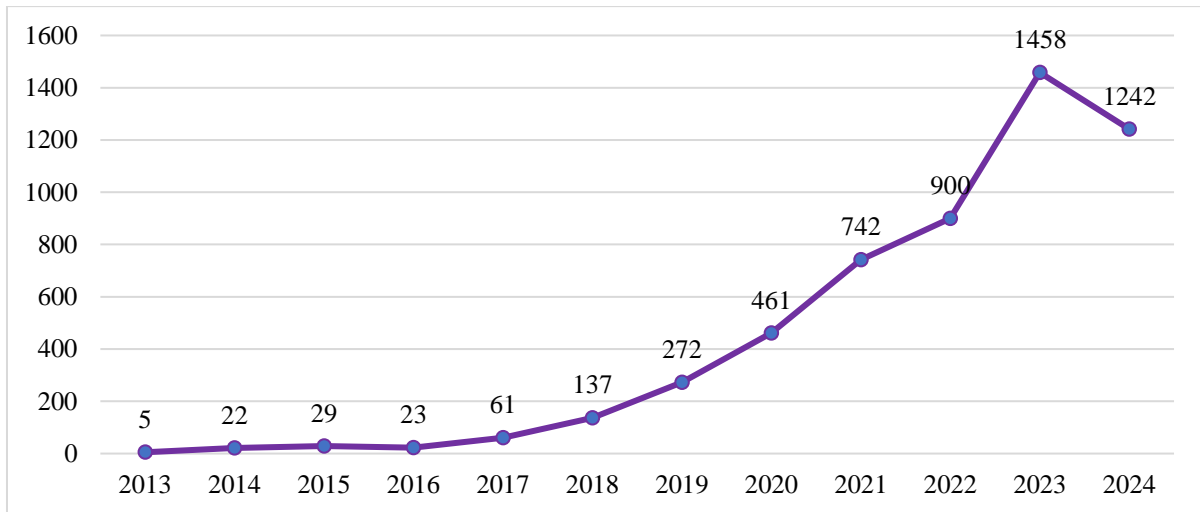
4. táblázat. A megjelent tanulmányok tudományterületi megoszlása az első hat kategóriában (forrás: Scopus)

Szakterület	Megjelent tanulmányok aránya (%)		
	Etika/erkölcs és felelősség	Etika/erkölcs és innováció	Etika/erkölcs és felelősség és innováció
Orvostudomány	24,3%	17,9 %	15,8 %
Társadalomtudományok	22,9 %	20,3 %	20,1 %
Ápolástan	10,6 %	6,2 %	6,5 %
Művészetek és bölcsészettudományok	10,4 %	7,8 %	7,2 %
Üzlet, menedzsment és számvitel	7,5 %	10,6 %	10,9 %
Informatikai tudományok	3,5 %	7,5 %	8,7 %

A három kulcsszókapcsolathoz tartozó további összehasonlításokat, valamint vizsgálati területeket az I. melléklet 3-6. táblázata, valamint az I. melléklet 1. és 2. ábrája mutatja be.

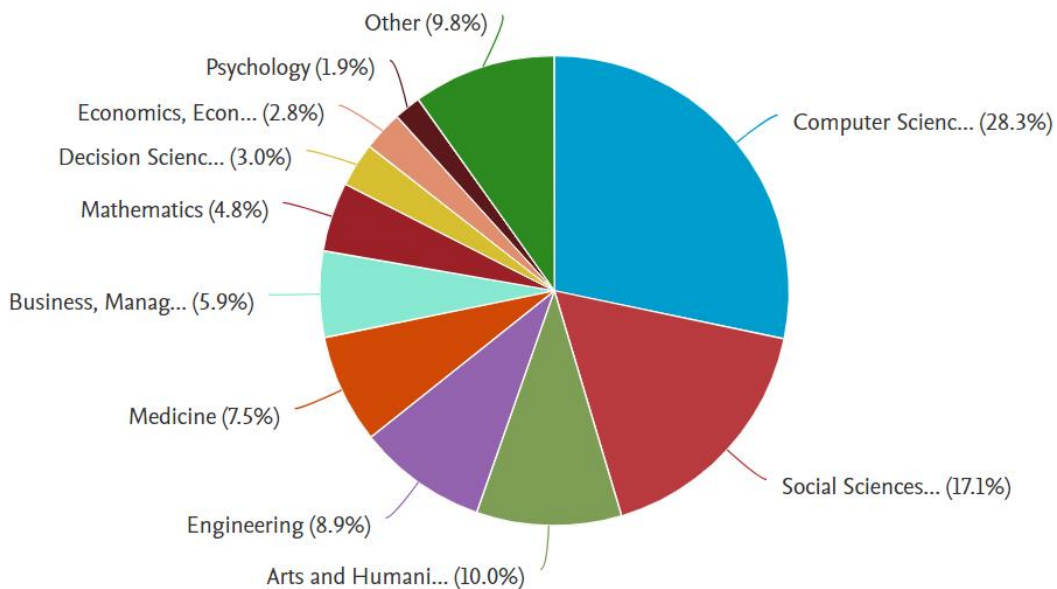
Megvizsgáltam továbbá egy másik nagyon fontos kategóriát a kulcsszavas keresés során, az 'etika/erkölcs és az AI' kapcsolatát ("*ethics*" OR "*moral*" AND "*artificial intelligence*" OR "*AI*"). 1996-tól indulva egészen 2024. június 30-ig összesen 7900 publikáció jelent meg, azonban ki kell emelni, hogy ez a publikációs szám 2024-ben már 1242-nél tart, amely adat már megközelítette a 2023-as éves összesített publikációs számot. A keresési vizsgálatból

továbbá kiderül, hogy 2013-ig szinte alig jelent meg tanulmány ebben a szókapcsolati vizsgálatban. A növekvő számadatokat a 9. ábra mutatja be.



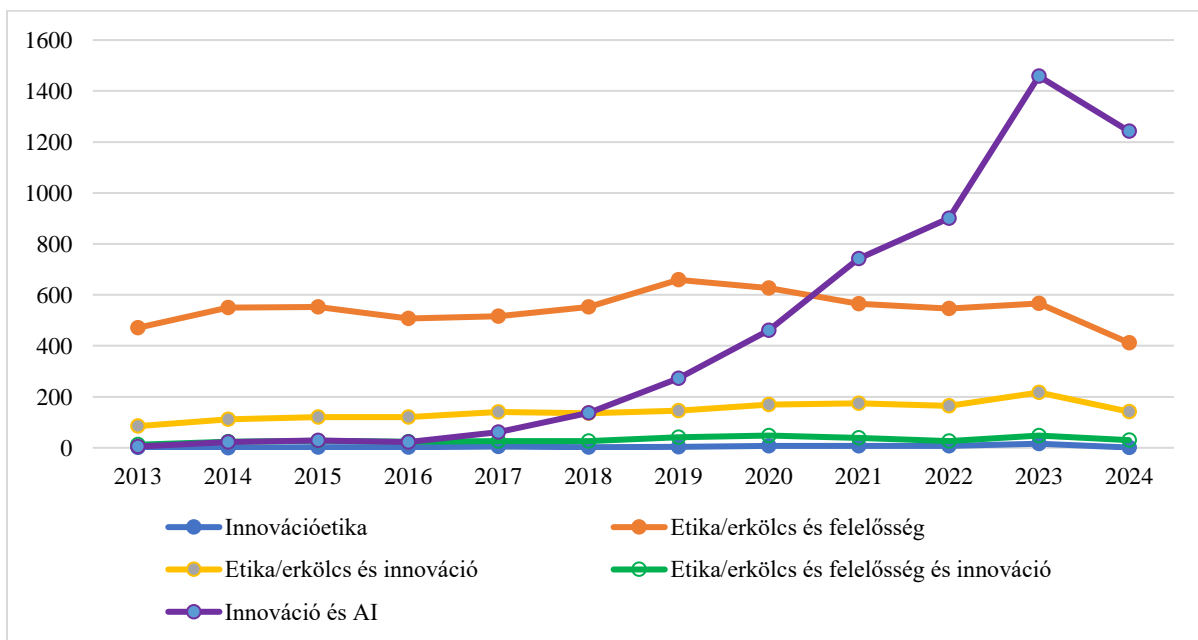
9. ábra. Az innováció és az AI kulcsszavak előfordulása (db) (forrás: Scopus)

A tudományterületi besorolásnál az 'innováció és AI' kulcsszavak kapcsolatánál hangsúlyosan megjelenik az informatikai tudományok szakterülete, amely a publikációknak majdnem a 30%-át képviseli, összesen 2919 db ilyen publikáció jelent meg a vizsgálati dátumig, és az is látható, hogy az előző keresésekben domináns orvostudomány csak az 5. helyen szerepel 7,5%-kal (10. ábra).



10. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és innováció) (forrás: Scopus)

Ha összesítve vizsgáljuk az öt kategóriát (11. ábra), akkor megállapíthatjuk, hogy jelentős növekedés figyelhető meg a publikációk számában 2017-től, amely 2023-ra éri el a csúcstot. Ez a trend arra utalhat, hogy az innovációval és az etikával kapcsolatos kérdések iránti érdeklődés erősen megnőtt az utóbbi években, ami feltehetően összefügg a technológiai fejlesztések felgyorsulásával és az ezekkel kapcsolatos etikai dilemmák növekedésével. Az 'etika/erkölcs és felelősség' területen a publikációk száma viszonylag stabil maradt 2013 és 2022 között, 2023-ban azonban egy jelentős csökkenés figyelhető meg. Ez a visszaesés jelezheti, hogy az etika és felelősség kérdéseit más, újabb területek háttérbe szorították. A COVID-19 időszakában (2020-2021) az 'innováció és AI' témakörben figyelhető meg a legnagyobb növekedés. Ez azt mutatja, hogy a világvárvány alatt jelentősen megnőtt az érdeklődés az AI-alapú innovációk etikai kérdéseiről. 2023-ra (sőt a 2024-es vizsgálati dátum időpontjához nézve) az összes témakörre vonatkozó találatok száma növekedést mutat, de nem olyan kiemelkedően, mint az AI-hoz kapcsolódó publikációk esetében. Ez az etikai kérdések növekvő fontosságára utal az innováció és mesterséges intelligencia diskurzusában.



11. ábra. A publikációk megoszlása évek és kategória szerinti bontásban (db) 2013-2024 (saját szerkesztés) (forrás: Scopus)

Végül pedig az összesített szakterületi táblázatából (5. táblázat) láthatóvá válik, hogy társadalomtudományi területen készült publikációk emelkednek ki arányaiban mind az 5 kategóriában, ez minden esetben 1. vagy 2. helyen szerepel. Míg az orvostudomány az 1. vagy a 2. helyet foglalja el az etika/erkölcs és felelősség és innováció hármassal minden kategóriájában, addig ez az innovációetika csak 3. helyen szerepel, sőt az innováció és AI kategóriában csak az 5. helyet foglalja el. Az innovációetika szakterületi publikációs arányainál második helyen

az üzlet, menedzsment és számvitel szerepel és az innováció és AI kategóriánál pedig az informatikai tudományok jelennek meg 1. helyen a táblázatban látható legmagasabb százalékkal.

5. táblázat. A megjelent tanulmányok tudományterületi megoszlása (forrás: Scopus)

Szakterület	Megjelent tanulmányok aránya (%)				
	Innovációetika	Etika/erkölcs és felelősség	Etika/erkölcs és innováció	Etika/erkölcs és felelősség és innováció	Innováció és AI
Orvostudomány	(3.) 14,7%	(1.) 24,3%	(2.) 17,9 %	(2.) 15,8 %	(5.) 7,5%
Társadalomtudományok	(1.) 16,7%	(2.) 22,9 %	(1.) 20,3 %	(1.) 20,1 %	(2.) 17,1%
Ápolástan	(10.) 2,7%	(3.) 10,6 %	(6.) 6,2 %	(6.) 6,5 %	-
Művészetek és bölcsészettudományok	(4.) 12,7%	(4.) 10,4 %	(4.) 7,8 %	(5.) 7,2 %	(3.) 10%
Üzlet, menedzsment és számvitel	(2.) 16%	(5.) 7,5 %	(3.) 10,6 %	(3.) 10,9 %	(6.) 5,9%
Informatikai tudományok	(6.) 6%	(6.) 3,5 %	(5.) 7,5 %	(4.) 8,7 %	(1.) 28,3%

* 5. helyen a 'Közgazdaságtan, ökonometria és pénzügyek' áll 8%-kal

**4. helyen 'Mérnöki tevékenység' áll 8,9%-kal

A bemutatott bibliometriai elemzések révén képet alkothattunk arról, hogy a dolgozatomban tárgyalt témák milyen gyakorisággal és kontextusban jelennek meg a tudományos publikációkban, alátámasztva témaválasztásom relevanciáját. A tudományos érdeklődés trendjeinek és azok változásainak a vizsgálata nem csupán arra világít rá, hogy melyek azok a tudományterületek, ahol aktív kutatás folyik, hanem arra is, hogy milyen irányban fejlődik a tudományos diskurzus. Az empirikus adatok integrálása hozzájárul ahhoz, hogy a dolgozat ne csupán elméleti, hanem gyakorlati alapokon nyugvó megállapításokat is tegyen, elősegítve annak a megértését, hogy a vizsgált fogalmak nem pusztán elméleti konstrukciók, hanem valós és aktív kutatási területek, amelyek jelentős szerepet játszanak a kortárs etikai témáinak kijelölésében.

A fenti adatok is rámutatnak arra, hogy az etikai dimenziók integrálása átformálja a globális tudománypolitikai paradigmát. Miközben az államok a technológiai fejlődés generálta átalakulással küzdenek, az etika szerepe az innovációs tevékenységek értékelésében minden eddiginél hangsúlyosabbá válik. A korábbiakban felvázolt dinamikus változások miatt szükséges, hogy alkalmazkodjunk az új helyzetekhez, valamint a globális biztonságot érintő stratégiák és eljárások adaptív újrahangolását és finomítását. Ahogy a mesterséges intelligencia

vagy a biotechnológia fejlődésének és paradigmaformáló erejének tanúi vagyunk, egyre nyilvánvalóbbá válik, hogy az innovációs tevékenységeket nem lehet az etikai, társadalmi és környezeti következmények figyelembevétele nélkül folytatni (Owen et al., 2013) – ebben kiemelt szerepe van egy felelősségközpontú szemléletnek, amelyre következő alfejezetben fókuszálunk majd.

4.5 A felelősségetika mint az innovációetika normatív keretrendszere

Az eddig leírtak alapján nyilvánvaló, hogy az innovációetikáról önálló (szub)diszciplináris területként nem beszélhetünk, a tudományos szakirodalomban szinte ismeretlen, hivatkozásai partikulárisak vagy marginálisak, keretei tisztázatlanok. Az alkalmazott etika egyik ágaként definiálható, azonban konceptuális határai elmosódtak, kialakulóban vannak – szaktetikaként még nem funkcionál. Amennyiben az innováció és a felelősség kontextusában érvényes megállításokat kívánunk tenni, szükségünk lesz egy olyan elméleti keretre, amely fogalmilag tisztázott elvrendszert kínál az elemzésre és a gyakorlati cselekvés normatív szabályait egyértelműen határozza meg. Az innovációetika számára lehetőséget kínálna, hogy betagozódjon a felelősségetika, mint alkalmazott etika ernyője alá – amennyiben a felelősségetikát komplementer normatív rendszerként kezeljük és elfogadjuk hatókörét az erkölcsi alapú döntések megalapozásában. Ennek legnagyobb akadályja, hogy magának a felelősségetikának is tisztázatlan a diszciplináris státusza.

Mint láthattuk az előző fejezetben, a felelősség egyre növekvő figyelmet kap a tudományos diskurzusban, például az orvosi, a technológiai, az üzleti vagy a környezeti etikában, azonban felelősségetika, mint önálló terület továbbra is kevésbé artikulálódott. Olyan gondolkodók, mint Max Weber (1919), Levinas (1974) vagy Hans Jonas (1984) munkái alapozták meg a felelősség etikai kontextusban való tárgyalását, és a felelősség kérdésköréről szóló vitákban megkerülhetetlenek a szakirodalomban (Gulyás, 2021). A hivatkozások ellenére, a felelősségetika mint területi etika, mégsem nyer önállóságot. Az előző fejezet módszertanára alapozva vizsgáltam meg a *felelősségetika* kifejezés előfordulását, kizárólagosan az *ethics of responsibility* kulcsszóra tovább szűkítve a szűrést, ami összesen 55 dokumentumot eredményezett az 1996-2004 közötti időszakban. Ezek csaknem kétharmad arányban oszlanak meg a társadalomtudományok (25 db – 26,3 %), a humántudományok (31 db – 32,6 %) illetve az orvostudományok (10 db – 10,5 %) között. A megjelent publikációk területi eloszlása itt nem mutat angolszász dominanciát – Németország, Dél-Afrika, Egyesült Királyság, Brazília és Franciaország az öt ország, amihez a találatok többsége köthető. Összességében kijelenthető, hogy a fogalom előfordulása sporadikus, a fellelt szakirodalmakat elemezve elmondható, hogy

kevés esetben történik önálló etikai területként utalás a felelősségetikára, a rendelkezésre álló adatok alapján a kortárs tudományosságban nagyon alacsonyan reflektált.

A probléma gyökerei oda vezetnek vissza, hogy hiányzik egy olyan átfogó elméleti munka, amely a felelősségetikát koherens és átfogó keretként határozná meg. A második fejezetben már beleütköztünk a felelősség definíciós problémájába, amit egy komplex, dinamikus fogalomként határoztam meg. Ez a komplexitás nem kedvez egy felelősséget fókuszába állító önálló rendszernek, mivel ahogy az előzőekben láttuk, még szűk területeken is nagyon eltérő megközelítéseket találunk a felelősség hatókörére vagy az ágenciára vonatkozóan – meg sem nyitva a vitát az olyan absztrakt kérdésekről, mint a szabadság. Jelen dolgozatban egy önálló felelősségetikai elmélet kidolgozását nem tudom felvállalni, azonban meghatározom azokat a kereteket, amelyekben a felelősségetikai elvek gyakorlati alkalmazása megtörténhet. Ez természetesen nem lehet független a felelősségetika kapcsán már korábban lefektetett alapoktól, így én elsősorban Hans Jonas gondolataira fogok támaszkodni, tekintetbe véve azok alkalmazhatóságának határait.

Ha a felelősségetika gyökereit keressük, elsőként Max Weber munkásságát kell megemlítenünk, aki a Politika mint hivatás című beszédében állítja szembe a lelkiismeret-etikának (Gesinnungsethik) és az általa felelősségetikának (Verantwortungsethik) nevezett szemléletet, amelyek eltérő erkölcsi cselekvéshez vezetnek. Weber értelmezésében a lelkiismeret-etikát, az erkölcsi elvekhez és eszmékhez való szigorú ragaszkodás jellemzi – anélkül, hogy a következményekre ugyanúgy hangsúlyt helyezne, a személyes meggyőződést tekinti elsődleges vezérlő elvnek. Ez a fajta cselekvés, amennyiben negatív következményekkel vagy hatásokkal jár „*a cselekvő azt gondolja, hogy ezért nem ő a felelős, hanem a világ, a többi ember ostobasága, netán Isten akarata, aki így teremtette őket*” (Weber, 2014, s.p.). A lelkiismeret-etika alapján, csak az elvek konzekvens követése iránt van felelőssége a cselekvőnek, így lényegében közömbös a tettei valós hatásaival szemben. A felelősségetika elveit követi, azonban számol a körülményekkel, felismeri a cselekedetek következményeit, mérlegeli a potenciális hatásokat, mert a „*cselekedetének (előrelátható) következményeit neki kell viselnie*” (Weber, 2014, s.p.). A politikai szerepet vállalók számára ez a megközelítés az ajánlott, mivel megköveteli a döntéshozóktól, hogy mérlegeljék a cselekedeteik szélesebbkörű társadalmi kihatásait. A jó politikai vezető rendelkezik arányérzékkel és kellő távolságtartással a megfelelő helyzetértékeléshez és racionális döntéshozatalhoz. A weberi koncepció egy pragmatikus megközelítés jelent: figyelembe kell vennünk tetteink hosszútávú

következményeit, tudnunk kell mérlegelni az erkölcsi problémák összetettségét és számolnunk kell a személyes felelősséggel – túllépve az absztrakt etikai elvekhez ragaszkodáson.

A francia filozófus Emmanuel Lévinas munkásságának középpontjában a felelősség egy Weberénél sokkal elvontabb fogalma áll. Legismertebb művében, a *Teljesség és végtelen-ben* alapozza meg etikáját, amely állítása szerint a *Másikkal* való találkozásból fakad. „*Az abszolút módon Más – a Másik. Nem vonható egybe az énnel*” – fogalmaz (Lévinas, 1999, p. 22), „*azon a módon, ahogyan a Más megjelenik, meghaladva a Más ideáját bennem, ténylegesen arcnak hívjuk*” (pp. 33). Ezzel az arccal való találkozás, nem pusztán észlelés, hanem a Másik feltárulkozása és ennek egy etikai imperatívusza van – a felelősségé. Ez a találkozás az alapja a felelősség tágabb kiterjesztésének is, a Másik általi megszólítás egyben aszimmetrikus kötelezettséget is jelent, a reciprocitás igénye nélkül. Az *Autrement qu'être ou au-delà de l'essence* (Másképp, mint a lét, avagy a lényegen túl) című könyvében a Másikkal szembeni felelősség nyitott koncepciója jelenik meg, amely azt feltétel nélkülivé és végtelenné teszi: az egyén a Másik *túszává* válik. A felelősség morális gyökerei a lét szubjektív alapját képezik, előtte vannak az emlékezetnek és a megértésnek:

„The responsibility for the other can not have begun in my commitment, in my decision. The unlimited responsibility in which I find myself comes from the hither side of my freedom, from a 'prior to every memory,' an 'ulterior to every accomplishment,' from the non-present par excellence, the non-original, the anarchical, prior to or beyond essence. The responsibility for the other is the locus in which is situated the null-site of subjectivity, where the question 'Where?' no longer holds”²²(Lévinas, 2006, p. 10).

A Lévinas által vázolt erkölcsi rend abszolút, a társadalom és a világ struktúráival szemben, egy azokat megelőző állapot, alapértékei nem kulturális konstrukciók. Az egyéni érdekek a Másiknak való alárendelése, az egyéni felelősség totalitása egy olyan radikális felelősségetika alapjait teremti meg, amelyben nem a tetteink milyenségéért vagy következményeinkért vagyunk felelősek, hanem felelősségünk valaki felé van – aki megszólított. A lévinasi gondolatok alapján, Lányi András szerint „*a felelősségetika tehát perszonalista, a személyek közötti viszonyt tekinti az erkölcsiség tárgyának, és kötelességeinket megszólítottságunkból, a konkrét kapcsolatból eredezteti*” (Lányi, 2012, p. 75).

²² „*A másikért való felelősség nem kezdődhet az én elköteleződésemben vagy döntésemben. Az a korlátlan felelősség, amelyben találok magam, a szabadságom határain túlról ered, egy olyan forrásból, amely »minden emlékezet előtt« létezik, és »minden teljesítményt meghalad«, a par excellence nem-jelenlétből, a nem-eredeti, az anarchikus, a lényeg előtt vagy azon túl létezőből fakad. A másikért való felelősség az a hely, ahol a szubjektív nullpontja található, ahol a »Hol?« kérdésének már nincs értelme.*” (saját fordítás)

A huszadik század második felének egyik legjelentősebb európai gondolkodója, Hans Jonas az 1979-ben megjelent *Das Prinzip Verantwortung* (A felelősség elve) című munkájában felhívja a figyelmet arra, hogy új kockázatok jelentek meg az emberiség számára: a technika fejlődése visszafordíthatatlannak tűnő változásokat okozva, közvetve az egész természetet és benne az emberi faj létét is fenyegeti. Jonas megállapítja, hogy erre az új helyzetre és kihívásra az etika nem reflektált – az ember technikai és társadalmi uralmának a kiterjedésével együtt ennek az új hatalomnak az erkölcsi alapjai és háttere szabályozatlan maradt (Szilágyi, 2021, p.41). A tudomány fejlődése aláasta azokat az alapokat, amelyekből korábban az etikai normáinkat levezettük, ezért egy általa *erkölcsi vákuumnak* nevezett helyzetben találjuk magunkat: „*Now we shiver in the nakedness of a nihilism in which near-omnipotence is paired with near-emptiness, greatest capacity with knowing least for what ends to use it*”²³ (Jonas, 1984, p. 23). Az idáig antropocentrikus etikai rendszerek a társadalmi élet területeire koncentráltak, az emberiség természetéhez való viszonya kívül esett a vizsgálódási körükön és a „*természet nem volt az emberi felelősség tárgya*” (Jonas, 2005, p. 28). Az ember és a természet viszonyának megváltozásával a felelősség új dimenziói jelentek meg, az új etikának a jelen cselekvéseinek hosszú távú következményeit is tárgyává kell tennie: korábban az ember jelenléte a világban adott volt és az etikai elméletek ebből indultak ki, most erkölcsi kötelezettségünk az emberiség fennmaradásának biztosítása lett. Etikai felelősségünk a jelenen túl a jövő generációkra is kiterjed és jelenbéli cselekedetünk potenciális jövőbeli hatásait is mérlegelnünk szükséges. Ennek összefoglalása a következő axióma:

*„Act so that the effects of your action are compatible with -the permanence of genuine human-life”: or expressed negatively: “Act so that the effects of your action are nor destructive of the future possibility of such life”; or simply: “Do not compromise the conditions for an indefinite continuation of humanity on earth”*²⁴ (Jonas, 1984, p. 11).

Jonas a felelősség két formáját különbözteti meg: az első a formális felelősség (formal responsibility), a már megtett cselekvésekre, tehát a múltra vonatkozik – cselekvésünkért és annak hatásaiért ki vállal felelősséget? Ez a felelősség kereteire vonatkozik, hogy milyen formális feltételek mellett tekinthető valaki felelősnek, „*minden formális cselekvés oksági*

²³ „*Most itt didergünk egy olyan nihilizmus meztelenségében, amelyben a közel-mindenhatóság a közel-ürességgel, a szinte korlátlan képességek azzal párosulnak, hogy a legkevésbé tudjuk, hogy milyen célokra használjuk őket.*” (saját fordítás).

²⁴ „*Cselekedj úgy, hogy cselekedeted hatásai összhangban legyenek az autentikus emberi élet fennmaradásával”; vagy negatívan kifejezve: „Cselekedj úgy, hogy cselekedeted hatásai ne legyenek rombolóak az ilyen élet jövőbeli lehetőségére nézve”; vagy egyszerűen: „Ne veszélyeztesd az emberiség határozatlan ideig tartó fennmaradásának feltételeit a Földön”*

terhe” (Jonas, 1984, p. 91) vagyis pusztán elszámoltathatósági tény. A szubsztantív felelősség (substantive responsibility) ellenben a jövőre irányul: ez a felelősség valamilyen meghatározott tárgyhoz/tartalomhoz kapcsolódik, elkötelezi magát a cselekvő olyan meghatározott tettek mellett, amelyek ezen tárgyakra vonatkoznak (megfogalmazható úgy is, mint „valamiért felelősnek lenni”) (Szilágyi, 2021, p.41). Ahogy Jonas fogalmaz „*I feel responsible, not in the first place for my conduct and its consequences but for the matter that has a claim on my acting*”²⁵ (Jonas, 1984, p. 92). Vagyis nem pusztán elszámoltatható vagyok valamiért, hanem úgy érzem, hogy felelős vagyok azért, hogyan cselekszem valamilyen helyzetben vagy teszek valamilyen ügyért, amely igényli az én beavatkozásom. A felelősség itt már nem reaktív, mint a formális felelősség esetében, hanem proaktív – ösztönzést jelent valami érdekében történő cselekvésre, amelyet helyesnek tartok. Ez a gondolat összekapcsolódik Lévinas elméletével, amennyiben a felelősség nem pusztán egy formális és absztrakt kötelesség, nem saját tetteink következményeire vonatkozik, hanem a rajtunk kívül álló (*a Másik*) az, akire irányul, annak a pusztán léte már tételezi a felelősség imperatívuszát. Jonas szerint először a *létezni kell* (ought-to-be) követelménye jelenik meg az objektum számára (aki a másik személy vagy létező, akinek vagy aminek a létezése erkölcsi követelményt támaszt), majd ezt követi a *tenni kell* (ought-to-do) a szubjektum (az „én” – aki rendelkezik a cselekvés hatalmával) számára, aki hatalma révén cselekvésre hivatott. Az objektum – a lévinasi értelemben vett Másik – létezése támaszt velünk szemben erkölcsi követelményt azáltal, hogy hatalmunktól vagy gondoskodásunktól függ, okot ad arra, hogy felelősséget érezzünk iránta:

*„The dependent in its immanent right becomes commanding, the power in its transitive causality becomes committed, and committed in the double sense of being objectively responsible for what is thus entrusted to it, and affectively engaged through the feeling that sides with it, namely, »feeling responsible«”*²⁶ (Jonas, 1984, p. 92).

Az emberek felelősségi viszonyba a formális felelősség esetében természetes módon kerülnek, a maguk elhatározásából, egyfajta szerződéses kapcsolat révén. A szubsztantív felelősség esetében az egyének és a tárgyak azonban függetlenül az akaratuktól válnak érintetté, egy Jonas által „*természetes és nem kölcsönös viszonyként*” jellemezve (Jonas, 1984, pp. 90-94, Szilágyi, 2022, p. 118).

²⁵ „*tehát nem elsősorban a magatartásomért és annak következményeiért érzem magam felelősnek, hanem azért a dologért, amely igényt tart a cselekvésemre.*” (saját fordítás)

²⁶ „*A (tőlünk) függő a maga immanens jogában parancsolóvá válik, a hatalom a maga tranzitív oksági viszonyában elkötelezetté válik, mégpedig kettős értelemben: objektíve felelős azért, ami így rá van bízva, és érzelmileg elköteleződik, azaz »felelősnek érzi magát«”* (saját fordítás)

Jonas megkísérel egy komplex paradigmaváltó etikát teremteni a technológiai társadalom számára, és fontos szerepet vállalt abban, hogy a felelősség kérdésköre újra hangsúlyt kapjon a jelen etikai diskurzusaiban. Az általa bevezetett új etikai elv, a *felelősség kötelezettsége* (imperative of responsibility) az emberi cselekvés minden területén egy óvatos és előrelátó megközelítést követel meg, az élet és a környezet fenntarthatóságának prioritását hirdetve.

Jonas igyekszik viszonyítási pontokat kínálni ebben, egy önálló, általa a *félelem heurisztikájának* (The Heuristics of Fear) nevezett döntéshozatali módszert mutatva be. Az általában bénító hatásúnak tartott félelem, a fenyegetettség előérzete, a lehetséges veszélyek és potenciálisan káros események előrejelzésének eszközeként, egyfajta iránytűként kerül felhasználásra. „*I contend that we need today an imaginative-anticipatory heuristics of fear to lead us to the discovery of the duties, even the ethical principles, with which to meet the challenge of coming events*”²⁷ – állítja (Jonas, 1976, p. 87). Jonas szerint a technológiai fejlődés következményeinek beláthatatlansága és potenciális negatív hatása miatt, az etikai mérlegelésben a negatív előrejelzéseknek elsőbbséget kell élvezniük „*stated primitively, that the prophecy of doom is to be given greater heed than the prophecy of bliss*”²⁸ (Jonas, 1976, p. 89). Azt, hogy ezt az elvet kövessük három fő okot jelöl meg, egyfelől a technológia fejlődésének üteme sokkal gyorsabb, mint a természetes evolúció üteme, ezért lényegében lemond a „biztonsági játékról”; másfelől a technológiai fejlődés gyorsasága nem hagy időt az önkorrekcióna és ennek amúgy sem lehetnének biztosak a kimenetelében, és harmadrészt a beindított folyamatok bármikor önállósodhatnak és kikerülhetnek az ellenőrzésünk alól. A korábbi etikák sokszor naiv optimizmusával szemben ezért kell a technológiai korszak etikájának a félelem előrejelző szerepére épülnie.

Jonas gondolkodásában a felelősség fogalma összekapcsolódik az egyén világra gyakorolt hatásával és rámutat, hogy ennek a hatóképességnek a felelősséggel összhangban kell lennie (Tóth, 2007, p. 55). Maga a felelősség egy aszimmetrikus viszonyt feltételez a hatást gyakorló és a hatás által érintettek között. Hans Jonas felelősségértelmezése az ember-ember viszonyokról, elmozdul az ember-környezet viszony felé, és a természetet is az emberi felelősség tárgyává teszi: mindenkinek akként kell cselekednie, hogy mérlegeli azt, hogy

²⁷ „Azt állítom, hogy ma szükségünk van a félelem egy képzelőerővel bíró-előrelátó heurisztikájára, hogy elvezessen minket azon kötelességek, sőt etikai alapelvek felfedezéséhez, amelyekkel szembenézhetünk az eljövendő események kihívásaival.” (saját fordítás)

²⁸ „egyszerűen megfogalmazva: a végítéletéről szóló próféciaúrnak nagyobb figyelmet kell szentelni, mint az üdvösségről szólóknak.” (saját fordítás)

cselekvése fenyegeti-e az élet jövőbeli fennmaradását – a felelősség ekként egyfajta „gondoskodó irányultsággá” válik (Szilágyi, 2022, p. 117). Kiemeli, hogy a kihívások globálisak, a technológiai forradalom globális kiterjesztése jelentős környezeti és társadalmi problémákat generálhat olyan fejlődő régiókban, amelyek nincsenek felkészülve a fejlődés potenciális veszélyeire. Ezért szükséges a globális politikában egy konstruktív, a helyi adottságokkal és a környezettel egyensúlyban lévő modell kialakítása: az emberi cselekvés következményeinek mérlegelni kell világméretű hatásait is. Jonas rámutat arra, hogy egy olyan kollektív gyakorlat kialakítására van szükség, ami összhangban van ezzel a jövőorientált új etikával, amely ragaszkodik a jövő nemzedékek iránti felelősség racionális követelményéhez (Szilágyi, 2022, pp. 41-43).

4.6 Az innovációs folyamatok felelősségetikai irányelvei

A felelősségetika különböző filozófiai megközelítéseinek áttekintése alapján látható, hogy a maga kiforratlanságában is megbízható elméleti alapot szolgáltat a kortárs etikai kihívások kezelésére az innovációetika kontextusában.

Ha a felelősségetikai alapelveket az innovációs folyamatra kívánjuk alkalmazni, négy konceptuális pillért tudunk meghatározni:

- [1] a *hosszú távú következmények figyelembevétele*, ami magában foglalja a közvetlen és közvetett hatások komplex hálózatának feltérképezését, különös tekintettel az időben elhúzódó, kumulatív és emergens jelenségekre;
- [2] az *elővigyázatosság elve* a fejlesztési folyamatokban rejlő inherens bizonytalanság elfogadása és felkészülés a potenciális negatív következményekre, illetve egy proaktív kockázatmenedzselési stratégia kialakítása;
- [3] a *globális hatások figyelembevétele* egy kiterjesztett felelősségi kört jelent, a bolygó egészét magába foglalva, potenciálisan csak lokális következményeket is beemelve az emberi cselekvés hatásainak mérlegelésébe;
- [4] a *jövő generációk iránti felelősség* a cselekvés etikai értékelésének kiterjesztése a még meg nem született generációk feltételezett érdekeire és jogaira, a jelen normái és jövőbeli generációk anticipált értékrendje közötti egyensúly megteremtésére törekedve.

A fenti négy pillér együttesen alkotják azt a normatív keretrendszert, amely szükséges az innovációs folyamatok felelősségteljes irányításához. A következő fejezetekben ezen keret gyakorlati alkalmazásának felvázolására teszek kísérletet egy esettanulmányon keresztül, amely a mesterséges intelligenciát állítja a fókuszába. A témaválasztásom kapcsán idézném a

fenti bibliometriai kutatást, amiből kitűnik, hogy az AI fejlesztésével és alkalmazásával kapcsolatos etikai kérdések az utóbbi években kiemelt figyelmet kaptak a tudományos diskurzusban. Láthattuk, hogy milyen mértékben növekedett meg a tudományos érdeklődés az elmúlt években, olyannyira, hogy lényegében napjaink legsokoldalúbban kutatott etikai kérdései közé emelkedett a mesterséges intelligencia fejlesztése és alkalmazása. További indokom volt, hogy a mesterséges intelligencia egyedülálló kihívások elé állítja az etikát, mivel gyors fejlődésének következményei talán a technológiai innováció más területeinél is kevésbé beláthatók, elég csak a döntéshozatal vagy a morális ágens kérdéskörére gondolni, amelynek közvetlen felelősségetikai implikációi vannak. Ki viseli a felelősséget egy AI-alapú rendszer cselekedeteiért? Meddig terjed a fejlesztők felelőssége és hol kezdődik magának a rendszernek az ágenciája, ha egyáltalán beszélhetünk ilyenről? Az AI olyan globális kihívásokat jelent, amelyek egyaránt hatnak a jövő generációira és a bolygó egészére is – az orvosdiagnosztikai eszközöktől az önvezető autókig közvetlenül érintik az emberi élet számos aspektusát.

Mindezek alapján úgy vélem, hogy a mesterséges intelligencia területe ideális terep a felelősségetikai elvek gyakorlati alkalmazhatóságának vizsgálatára, kiemelten a saját felelősségetikai alapokon nyugvó értékelési modellem tesztelésére.

V. A mesterséges intelligencia etikai és társadalmi kihívásai

Az esettanulmányok integráns részét képezik az alkalmazott etikai praxisnak, mert lehetővé teszik, hogy etikai elveket és elméleteket valós helyzetekben vizsgáljunk. Ilyen példa Beauchamp és Childress már idézett *Principles of Biomedical Ethics* alapművének újabb kiadása, amely számos esetleírást és esettanulmányt tartalmaz etikai problémák szemléltetésére (Beauchamp & Childress, 2012). Illetve kiemelném az Egyesült Államokban a National Institutes of Health-t, ami *Annual Review of Ethics Case Studies* címen tesz elérhetővé tematikus esettanulmányokat, amelyek komplex etikai problémákat bemutatva segítik a kutatókat az ilyen jellegű kihívások kezelésében szükséges készségek fejlesztésében (NIH, 2024). A következőkben egy ilyen esettanulmányon keresztül mutatom be a felelősségetikai elvek működését a mesterséges intelligencia fejlesztésében.

5.1 Felelős innováció és etikai megfontolások az AI-fejlesztésben

Merton azt állítja, hogy a tudományos közösséget az „innováció ethosza” dinamizálja, ez ösztönzi a tudósokat, hogy úttörő fejlesztési tevékenységekbe kapcsolódjanak be és új eredményeket hozzanak létre (Merton, 1957). A tudományos kutatásba vetett közbizalom fenntartása, valamint a kutatói közösségek, intézmények és a tudásbázisok hosszú távú fennmaradásának biztosítása érdekében elengedhetetlen valamilyen közös etikai alaphoz megteremtése. Paula Stephan a kutatási források elosztását és a kutatási környezet változását vizsgáló *How Economics Shapes Science* (Stephan, 2012) című könyvében azt javasolja, hogy mind az akadémiai, mind az ipari intézmények kövessenek egy hasonló stratégiát az innovációs tevékenységek ösztönzése terén. A kutatókat ne csak elismerjék és jutalmazzák innovatív tevékenységükért ösztöndíjjal, előléptetések, kitüntetések és szakmai presztízs elnyerése révén, hanem tudományos pályafutásuk során az innováció értékének megbecsülésére is neveljék őket, új szakmai normákat állítva. A produktív, kreatív és egyben elszámoltatható tudományos fejlődés előmozdítása érdekében jelentős etikai és jogi változásokra, újfajta ösztönzőkre és szélesebb körű átalakulásra lehet szükség a tudományos kultúrában. A tudomány értéksemleges ideálja egyre inkább háttérbe szorul és az értékek mindinkább jelentős szerepet kapnak a kutatási témák kiválasztásától, az információs csatornák kiválasztásáig. Ezeknek az értékeknek valamilyen módon tükrözniük kell a társadalom etikai preferenciáit és kellő transzparenciával kell megjeleníteniük mindenki számára (Elliott, 2017). Szükségessé válik, hogy a kutatással foglalkozó szervezetek etikai kódexeket vezessenek be, tudományetikai képzést biztosítsanak, és olyan szabályokat érvényesítsenek, amelyek világosan körvonalazzák a szervezet

alkalmazottakkal szembeni elvárásait és az e normák megsértésének lehetséges következményeit.

A felelős innováció kultúrájának előmozdítása kulcsfontosságú elem a mesterséges intelligencia adaptív kontrolljának kiépítésében. Ez magában foglalja egy, a szingularitás problematikáját fókuszban tartó proaktív és elővigyázatos gondolkodásmód ösztönzését a mesterséges intelligencia fejlesztésével foglalkozók körében éppúgy, mint etikai megfontolások integrálásának előmozdítását az innovációs folyamat során (Stilgoe et al., 2013), továbbá az átláthatóság, az elszámoltathatóság és a nyilvánosság bevonása iránti elkötelezettséget a fejlesztői-kutatói oldalon (Szilágyi, 2024a, pp. 142-146).

Floridi és munkatársai, az Atomium-EISMD a Good AI Society alapjainak lefektetése céljából létrehozott AI4People kezdeményezésének keretében vizsgálták a mesterséges intelligencia fő társadalmi lehetőségeit és kockázatait, és öt etikai alapelvet fogalmaztak meg, amelyeknek a technológia fejlesztését és alkalmazását meg kell alapozniuk. A fejlesztés során elsődleges alapelvként figyelembe kell vennünk, hogy a mesterséges intelligencia [1] jótékony és [2] ne ártalmas legyen (beneficence és non-maleficence), elő kell segítenie és nem korlátoznia az [3] emberi autonómiát (autonomy), és végül mesterséges intelligencia kapcsán megjelenik még az [4] igazságosság (justice) kritériuma is, vagyis biztosítva kell legyen a technológiát fejlesztő és alkalmazó emberek és szervezetek – felelősségre vonásának lehetősége – ami viszont megköveteli annak megértését, hogy miért következik be az adott eredmény. Emiatt mindez kiegészül egy új alapelvvel, a [5] megmagyarázhatósággal (explicability), amely két szempontot foglal magában: az érthetőséget (intelligibility) a kifejezés episztemológiai értelmében, amely a *hogyan működik?* kérdést válaszolja meg, és az elszámoltathatóságot (accountability) etikai értelemben, amely arra a kérdésre ad választ, hogy *ki felelős azért, ahogyan működik?* (Floridi et al., 2018, pp. 699-700).

Mindezen elvek gyakorlati alkalmazása csak akkor működhet, ha megteremtjük a kutatói-fejlesztői oldalon hajlandóságot arra, hogy az AI-rendszereket az új felismerések és aggályok hatására módosítsák: emiatt az etikai megfontolások integrálása az innovációs folyamatokba alapvető fontosságú annak biztosításához, hogy a mesterséges intelligenciát használó rendszerek fejlesztése felelősségteljes módon menjen végbe, figyelembe véve az AI-technológiák szélesebb körű társadalmi és etikai következményeit.

5.1.1 A mesterséges intelligencia etikai szabályozásának fejlődése

A mesterséges intelligencia fejlődése és térnyerése napjaink egyik olyan meghatározó jelensége, amely egyszerre hordoz magában be nem látható lehetőségeket és komoly

kockázatokat. A szabályozási kísérletek többsége is kockázatelvű, ami azt jelenti, hogy a figyelmet azokra a területekre fókuszálja, amelyek a legsérülékenyebbek, ahol a legnagyobb valószínűsége a potenciális kárnak, amelyek biztonsági szempontból a legsebezhetőbbek. Ha összevetjük ezeket a kísérleteket nagyon hasonló beavatkozási területeket találunk, ami egyfajta globális konszenzus kialakulása felé mutat. Mindezek mellett ki kell emelni azt, hogy a kockázatelvű megközelítés, habár hasznos eszköz a szabályozási prioritások meghatározásában, önmagában nem helyettesítheti az etikai szempontok lehatárolását, ahogy a mesterséges intelligencia felelősségteljes fejlesztése és alkalmazása sem pusztán etikai kérdés, hanem fontos jogi és szabályozási kihívás.

Az elmúlt években számos irányelv látott napvilágot a felelősségteljes innováció, a felelős AI-fejlesztés kapcsán, ám ezek többségében jogilag nem kötelező érvényű ajánlások voltak csak, így betartásuk a vállalatok önkéntes elkötelezettségén és vállalásain múlt. A 2024-es év hozott jelentős áttörést, amikor az Európai Unióban érvénybe lépett (2024. augusztus elsején) az AI Act-ként ismertté vált mesterséges intelligenciáról szóló törvényerejű rendelet, amely minden tagállamban kötelező és közvetlenül alkalmazandó. A jövőben várhatóan újabb országok vezetnek be az eddigieknél szigorúbb szabályozásokat, így például Kanada vagy az Egyesült államok is több törvénytervezetet készített elő az AI szabályozására és a fejlesztéssel járó kockázatok csökkentésére (Habuka & Socol de la Osa, 2024). A következőkben az elmúlt egy évtized legfontosabb szakmai dokumentumait tekintem át, azzal a szándékkal, hogy ezek alapján meghatározzam a legfontosabb kulcsponthoz és vizsgálandó területeket egy alkalmazott etikai elemzés elvégzéséhez a mesterséges intelligencia tárgyában. A célom az, hogy kritikai szemszögből vizsgáljam a kulcsdokumentumokat, összehasonlító elemzését nyújtva a fő fókuszterületeknek és alapelveknek.

5.1.2 Nemzetközi irányelvek és szabályozási törekvések az AI területén

A 2017-ben a neves technológiai szakemberek által alapított nonprofit szervezet, a Future of Life Institute szervezte az Asilomar Conference on Beneficial AI találkozót, ahol a résztvevő tudósok 23 irányelvet fogadtak el a mesterséges intelligencia kutatására és fejlesztésre vonatkozóan. Az irányelvek kidolgozásában a konferencián résztvevő technológiai szakértők, AI-kutatók, robotikusok és jogi szakértők egyaránt részt vettek, és a kiadás óta több ezer szakember és nagy technológiai cég is csatlakozott. Az úgynevezett Asilomar-alapelvek (Future of Life Institute, 2017) a mesterséges intelligencia kutatásának, fejlesztésének és alkalmazásának az első nagyobb hatású önkéntes szabályozási kísérleteként ismertek. Az egyik alfejezet az Ethics and values címet viseli és 13 pontban fogalmazza meg az elveket,

amelyek közvetlen etikai kérdéseket érintenek, mit például a fejlesztések összehangolása az emberi értékekkel és értékrendekkel, a működés átláthatóságának biztosítása, a személyes adatok védelme vagy az emberi jólét előmozdítása. Rendkívül izgalmas felvetés fogalmazódik meg a 8-as alapelvben (Judicial Transparency) az igazságszolgáltatási döntéshozatali folyamatokban való részvétel kapcsán: „*Any involvement by an autonomous system in judicial decision-making should provide a satisfactory explanation auditable by a competent human authority*”²⁹(Asilomar AI Principles, 2017, s.p.). Ez a jövőre vonatkozóan egy komplex jogi és etikai problémakört nyit meg csakúgy, mint az autonóm fegyverrendszerekre vonatkozó 18. ajánlás, amely a fegyverkezési versenyben elővigyázatosságra int az AI alkalmazása kapcsán.

2018-ban indult el Nagy-Britanniában az Alan Turing Intézet Public Policy Programja, amelynek a célja kutatások végzése és új eszközök kifejlesztése a kormányzat számára, hogy támogassa az adatintenzív technológiai innovációt. A program kiemelt jelentőséget tulajdonít az etikai és biztonsági protokollok kifejlesztésének, ennek érdekében bocsájtotta közre 2019-ben az Understanding Artificial Intelligence Ethics and Safety (Leslie, 2019) dokumentumot, amely egy útmutató a mesterséges intelligencián alapuló technológiák fejlesztői és alkalmazói számára a közszférában. Maga a dokumentum komplex bevezetést nyújt az alkalmazandó irányelvekbe és biztonsági protokollokba, számos elméleti meglátással gazdagítva az AI és az etika viszonyrendszerét kutatókat. A szöveg az első oldalán definiálja a vizsgálatának tárgyát: „*AI ethics is a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies*”³⁰ (Leslie, 2019, p. 3). E dokumentum veti fel elsőként, hogy szükség van egy olyan etikai platformra, amely a gyakorlatba is átülteti a felelős innováció kultúráját, figyelmet fordítva a fejlesztések során az etikus cselekvésre, amelyet négy alapelvben ragadott meg (ezek az úgynevezett FAST-track principles): a méltányosság (fairness) elve, amely azt jelenti, hogy az AI rendszereknek diszkriminációmentesnek kell lenniük, kerülniük kell az elfogultságot és a megkülönböztető bánásmódot; az elszámoltathatóság (accountability) elve értelmében az AI fejlesztőknek és az ezt üzemeltetőknek felelősséget kell vállalniuk a rendszerek működéséért és annak a következményeiért; harmadik elvként a fenntarthatóság (sustainability) jelenik meg, ami ebben a kontextusban azt jelenti, hogy a mesterséges intelligenciát használó fejlesztéseknek figyelembe kell venniük a környezeti és társadalmi hatásokat és hosszú távon fenntarthatónak

²⁹ „Az autonóm (AI) rendszerek igazságszolgáltatási döntéshozatalban való bármilyen részvételének kielégítő magyarázatot kell nyújtania, amelyet egy illetékes emberi hatóság ellenőrizhet.” (saját fordítás)

³⁰ „Az AI etikája olyan értékek-, elvek- és technikák összessége, amely a mesterséges intelligencia fejlesztésében és használatában alkalmazható – jó és rossz gyakorlatokat egyaránt vizsgáló – erkölcsi iránymutatást nyújtó szabályokkal foglalkozik.” (saját fordítás)

kell lenniük; végül negyedikként az átláthatóság (transparency) elve, amely az AI rendszerek működésének, döntéshozatali folyamatainak és potenciális hatásainak átláthatóságát és érthetőségét határozza meg. Az útmutató számos részletes elemzésen és gyakorlati példán keresztül mutatja be ezeknek az alapelveknek a működését, pontos leírást adva az alkalmazásukra.

Az Európai Bizottság 2018-ban egy mesterséges intelligencia szakértői csoportot hozott létre High-Level Expert Group on Artificial Intelligence (AI-HLEG) néven, hogy segítse a Bizottság munkáját és szakértői agyagokat szolgáltatson az AI-al kapcsolatos stratégiájának kialakításában. A szakértői csoport által 2019-ben közreadott Ethics Guidelines For Trustworthy AI tartalmazza azokat a kulcsfontosságú alapelveket, amelyek meghatározzák napjaink EU-s jogalkotását a mesterséges intelligenciát illetően (AI-HLEG, 2019). A dokumentum hangsúlyozza, hogy az AI rendszerek fejlesztése és alkalmazása során figyelembe kell venni az emberi autonómia tiszteletben tartásának (respect for human autonomy) elvét, ami biztosítja, hogy az emberek teljes és hatékony önrendelkezést tarthassanak fenn; a károkozás megelőzésének (prevention of harm) elve szerint a mesterséges intelligenciát használó rendszerek nem okozhatnak kárt vagy más módon hátrányt embereknek; a méltányosság elve (fairness) megköveteli, hogy a fejlesztése és az alkalmazás a költségek és a haszon egyenlő, igazságos elosztásának a jegyében történjék, és végül a megmagyarázhatóságnak (explicability) az etikai elvét, ami magában foglalja a fejlesztési folyamatok átláthatóságát, és az elvárást, hogy az AI képességeit és célját érthetően és nyíltan kommunikálják. A dokumentum rámutat, hogy foglalkozni kell ezen elvek közötti összefüggésekkel és potenciális feszültségekkel is az alkalmazás során, továbbá figyelmet kell fordítani a kiszolgáltatott társadalmi csoportokra és a hatalmi vagy információs aszimmetriával leírható helyzetekre (AI-HLEG, 2019). A szakértői csoport további hét, nem kötelező érvényű erkölcsi irányelveket alkotott meg, amelyek támogatják a mesterséges intelligencia etikus és megbízható fejlesztését és alkalmazását, tovább erősítve a technológiába vetett bizalmat: *„emberi cselekvőképesség és felügyelet; műszaki stabilitás és biztonság; a magánélet védelme és adatkezelés; átláthatóság; sokszínűség, a megkülönböztetés tilalma és méltányosság; társadalmi és környezeti jóllét; valamint elszámoltathatóság”* (Európai Parlament és a Tanács, 2024, p. 8).

Az Európai Parlament 2020. október 20-i „A mesterséges intelligencia, a robotika és a kapcsolódó technológiák etikai szempontjainak kerete” – című állásfoglalása, az etikai megfelelés tanúsítása érdekében:

„javasolja, hogy az uniós szintű koordináció keretében dolgozzanak ki közös kritériumokat és egy kérelmezési eljárást egy etikai megfelelést igazoló európai tanúsítvány odaítéléséhez kapcsolódóan (...) úgy véli, hogy egy ilyen etikai megfelelést igazoló európai tanúsítvány a mesterségesintelligencia-ökoszisztémák ellátási láncának egészében előmozdítaná a beépített etikát; javasolja ezért, hogy a magas kockázatú technológiák esetében ez a tanúsítvány legyen a pályázatok elfogadhatóságának kötelező előfeltétele a mesterséges intelligenciára, robotikára és kapcsolódó technológiákra vonatkozó közbeszerzési eljárásokban”(Európai Parlament, 2020, 135-136. szakasz).

Az állásfoglalás kiemeli, hogy az Európai Uniónak előremutató szerepet kell betöltenie a mesterséges intelligencia etikus fejlesztésében és használatában, olyan módon, hogy szerepet vállal a globális sztemderdek kialakításában, amelyek biztosítják az emberi ellenőrzést, az átláthatóságot, az elszámoltathatóságot, a diszkriminációmentességet, a magánélet védelmét és a társadalmi jóllétet. Az etikai normák kidolgozásában és terjesztésében nagy jelentősége van a nemzetközi együttműködésnek, folyamatos párbeszédet folytatva az eltérő szabályozási modelleket követő országokkal, az etikai szabályozások globális elfogadása érdekében (Európai Parlament, 2020, 137-144. szakasz).

Az Egyesült Nemzetek Nevelésügyi, Tudományos és Kulturális Szervezete (UNESCO) 2021-ben fogadta el a Recommendation on the Ethics of Artificial Intelligence címet viselő dokumentumot, amelyben azt ajánlja a tagállamoknak, hogy önkéntesen alkalmazzák az abban foglalt rendelkezéseket és tegyék meg a megfelelő jogalkotási és harmonizációs intézkedéseket, hogy joghatóságukon belül érvényesíteni tudják ezeket. A dokumentum bevezetőjében felhívják a figyelmet az összes AI technológiák fejlesztésében és alkalmazásában érintett fél bevonására a párbeszédbe és az etikai alapelvek figyelembevételére, mivel hangsúlyozzák: a mesterséges intelligenciát használó rendszerek az életciklusuk minden szakaszában jelentős etikai kérdéseket vetnek fel. Az ajánlás az AI rendszerekhez így közelít *„as systems which have the capacity to process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control”*³¹ (UNESCO, 2021, p. 4), és részletesen felsorolja, hogy milyen etikai kihívásokkal szembesítenek minket a digitális szakadéktól a biztonságpolitikai és a jogállamiságot érintő kérdésekig. Az ajánlás rendkívül átfogó és tíz olyan alapelvet fogalmaz meg, amelyek iránymutatásul szolgálhatnak mind a kormányzati, mind a fejlesztői oldal számára. Ezek a következők: az AI használatának *biztonságosnak* kell lennie és nem okozhat kárt sem emberi

³¹ „Az ajánlás a mesterséges intelligencia rendszereket olyan rendszerekként közelíti meg, amelyek képesek adatokat és információkat feldolgozni intelligens működést feltételező módon, és jellemzően birtokolják az érvelés, a tanulás és az észlelés előrejelzés, tervezés vagy ellenőrzés aspektusait.” (saját fordítás)

életben, sem a környezetben, mindezek mellett törekedni kell arra, hogy az AI elősegítse a fenntartható, magánélet védelmét segítő és adatvédelem-központú fejlesztést és használatot. A mesterséges intelligencia használatának elő kell mozdítania a *méltányosságot és a diszkriminációmentességet* elvét a nemzetközi joggal összhangban, biztosítva azt, hogy az AI előnyei mindenki számára elérhetők és hozzáférhetők legyenek. Prioritásként kell kezelni a *fenntarthatósági célokat*, és folyamatosan végezni az AI emberi, társadalmi, kulturális, gazdasági és környezeti hatásainak mérését és értékelését, hogy pontos képet alkothassunk a fenntarthatóságra gyakorolt befolyásáról. Kiemelt jelentőségű a *magánélethez való jog és a személyes adatok védelme*, amelyet erősíteni és támogatni kell megfelelő adatvédelmi rendszerek és ellenőrzési mechanizmusok kiépítése révén. Biztosítani kell, hogy a fejlesztőknek és alkalmazóknak elszámoltathatónak kell lenniük az általuk létrehozott és használt rendszerekért, etikai és jogi felelősséggel, amely magán vagy jogi személyekhez rendelhető. Az AI rendszereknek *átláthatónak és megmagyarázhatónak* kell lenniük, az emberek teljes körű tájékoztatása és tájékozottsága az alapfeltétele az elszámoltathatóságnak és felelősségnek, ahogyan az átláthatóság a biztonság egy alappillére, lehetővé téve a nyilvános ellenőrzést a végfelhasználók számára érdemi válaszokat kínálva a működés és a hatás specifikumaira. A *felelősség és az elszámoltathatóság* kulcsfontosságú alapelvek: a mesterséges intelligencia által hozott döntésekért, a végrehajtott cselekedetekért és azok következményeiért mindig az emberek a felelősek, attól függően, hogy milyen szerepet töltek be az AI rendszer életciklusának különböző szakaszaiban (a fejlesztéstől a piaci alkalmazásig). Ennek megfelelően kell kialakítani a felügyeleti és hatásvizsgálati mechanizmusokat. A megmagyarázhatóság alapfeltétele az AI rendszerekkel kapcsolatos *ismeretek* birtoklása, a felhasználók széleskörű tájékoztatása és az esetleges hatások ismertetése, aminek érdekében a AI-al kapcsolatos ismereteket hozzáférhető oktatással és a nyilvánosság bevonásával kell előmozdítani. Végül a nemzeti és nemzetközi szabályozásoknak *adaptívnak és együttműködőnek* kell lenniük, tiszteletben tartva a nemzetközi jogszabályokat és a nemzeti szuverenitást egyaránt (UNESCO, 2021, pp.7-10). Az UNESCO az ajánlásában célul tűzi ki, hogy az AI fejlesztése és alkalmazása során a transzparencia, az elszámoltathatóság és az adatvédelem biztosításával, valamint az élet különböző területeire vonatkozó gyakorlatorientált szakpolitikai szempontok figyelembevételével, a mesterséges intelligencia előnyei érvényesüljenek, miközben a potenciális veszélyek minimalizálásra kerülnek.

Az OECD AI Policy Observatory és az OECD AI Expert Group on AI Futures munkacsoportjának kutatói adták közre 2023-ban az Initial policy considerations for generative

artificial intelligence munkaanyagot, ami annak ellenére, hogy konkrét etikai alapelveket nem fogalmaz meg, a generatív AI alkalmazása kapcsán számos potenciális kockázatot sorra vesz. A szerzők szerint a döntéshozói fórumoknak figyelembe kell venniük a generatív AI kiszámíthatatlan társadalmi és politikai hatásait, beleértve a munkaerőpiacra gyakorolt hatását, a szerzői jogi kérdéseket és az AI rendszerek manipulatív használatának kockázatait. Rámutatnak, hogy az AI hozzájárulhat a társadalmi előítélek kialakulásához és fenntartásához, hamis információk terjesztéséhez és tartós adatvédelmi és biztonsági kockázatokat vet fel az, hogy a technológia korábban soha nem látott adatmennyiségeket képes kezelni. A generatív AI széleskörű alkalmazása egyfajta technológiai függéshez vezethet, és a humán ágencia visszaszorulását idézheti elő, az emberi döntéshozatal és kontroll csökkenésével együtt, ami súlyos társadalmi, politikai és gazdasági következményekhez vezethet (Lorenz et al., 2023).

Az OECD 2019. május 22-én miniszteri szintű ülésén fogadta el a *The Recommendation on Artificial Intelligence* ajánlást – a mesterséges intelligenciára vonatkozó első kormányközi szabványt, amely 2023 novemberében lett felülvizsgálva és kiegészítve majd 2024-ben közreadva *Recommendation of the Council on Artificial Intelligence* címmel. Az ajánlás célja az innováció és a bizalom előmozdítása a mesterséges intelligencia területén azáltal, hogy elősegíti a megbízható mesterséges intelligencia felelősségteljes felügyeletét, biztosítva egyúttal az emberi jogok és a demokratikus értékek tiszteletben tartását. (OECD, 2024, p. 8)

A Hiroshima AI process-t 2023 májusában indították el a G7 országok (Kanada, Franciaország, Németország, Olaszország, Japán, Egyesült Királyság, Egyesült Államok és az EU) hiroshimai csúcstalálkozóján, felismerve egy globális AI szabályzási rendszer megalkotásának szükségességét. A nemzetközi szabályrendszer kialakítása 2023 decemberében zárult le és ekkor született a megállapodás az első nemzetközi ajánlásról, amelyet *Hiroshima Process International Guiding Principles for All AI Actors (AIP&CoC, 2023)* címen adtak közre. Az ajánlás 11 irányelvet fogalmaz meg, amelyek az AI életciklusának minden szereplőjére vonatkoznak, hangsúlyozva többek között a fejlesztéssel és az alkalmazással járó kockázatok azonosításának, értékelésének és kezelésének fontosságát, valamint a fejlesztők és más érdekelt felek közötti együttműködés és információmegosztás jelentőségét. Az ajánlás kiemeli az átláthatóságot, az elszámoltathatóságot és a biztonság központi szerepét, és ösztönzi a megbízható tartalomhitelesítési mechanizmusok alkalmazását, az AI által generált tartalmak azonosíthatósága érdekében. Az alapelvek arra is felhívják a figyelmet, hogy a fejlett AI-rendszereknek a globális kihívások (mint például az éghajlati válság) megoldását prioritásként kell kezelniük, továbbá tekintettel kell lenniük a személyes adatok és

a szellemi tulajdon védelmére. A fentiek alapján született meg a mesterséges intelligenciát használó rendszereket fejlesztő szervezetek számára a Hirosima-folyamat nemzetközi magatartási kódexe (Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems), amely útmutatást nyújt a 11 irányelv végrehajtásához, részletesen felsorolva azokat a lépéseket, amelyeket a fejlesztőknek be kell tartaniuk (AIP&CoC, 2023).

Az Európai Parlament és a Tanács (Eu) 2024/1689 Rendelete „a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról” (EU Artificial Intelligence Act – továbbiakban EU AI Act) 2024. augusztus 1-én lépett életbe és már az első szakaszaiban visszautal a 2019-es Ethics Guidelines For Trustworthy AI szakértői anyagra, mint etikai igazodási pontra, megismételve az abban lefektetett hét alapelvet, részletesen kifejtve és értelmezve azokat (Európai Parlament és a Tanács, 2024, p. 8, 27. szakasz). A rendelet lényegében az első átfogó jogi keretrendszer a mesterséges intelligenciával kapcsolatban, megszabva egyértelmű kötelezettségeket és követelményeket az AI-fejlesztők és felhasználók számára. A 167. szakaszban a nem magas kockázatú AI rendszerek kapcsán fogalmazódik meg az elvárás, hogy ösztönözni kell a szolgáltatókat önkéntes magatartási kódexek kialakítására, amelyeknek ki kell térniük például a környezeti fenntarthatóságra, az inkluzivitásra, a sokszínűségre, az érdekelt felek bevonására és a fejlesztői csapatok diverzitására. *„A kódexek hatékonyságát egyértelmű célkitűzésekkel és teljesítménymutatókkal kell biztosítani, Hatékonyságuk biztosításának érdekében az önkéntes magatartási kódexeknek egyértelmű célkitűzéseken és az e célkitűzések elérésének mérésére szolgáló fő teljesítménymutatókon kell alapulniuk”* (p. 41, 165. szakasz).

5.1.3 Összehasonlító elemzés: AI etikai irányelvek és szabályozások

Ahogy az előzőekben láthattuk, a különböző nemzetközi szervezetek az elmúlt évek során számos iránymutatást és ajánlást dolgoztak ki a mesterséges intelligencia etikai vonatkozásainak szabályozására. Az alábbi táblázat ezeket foglalja össze, felvázolva az egyes dokumentumok fő irányelveit, koncepcióit és fókuszterületeit. Ez az összehasonlító táblázat lehetővé teszi, hogy nyomon kövessük a mesterséges intelligencia etikai kereteinek fejlődését, feltárjuk a közös mintázatokat és a megkülönböztető jegyeket, valamint értékeljük, hogy ezek a szövegek miként kezelték az AI területén bekövetkezett fejlődést és az ezzel kapcsolatban megfogalmazódó új kérdéseket.

6. táblázat. Legfontosabb alapelvek és fókuszterületek (saját szerkesztés)

Dokumentum	Legfontosabb alapelvek	Fókuszterületek
Asilomar Principles (2017)	etika, átláthatóság, a magánélet védelme, emberi jólét, felelősség, emberi ellenőrzés	értékek egyeztetése, jogi kérdések, közös jólét, emberi értékek, biztonság, magánélet, mesterséges intelligencia fegyverkezési verseny, közjó
Understanding AI Ethics and Safety (2019)	méltányosság, elszámoltathatóság, fenntarthatóság, átláthatóság	diszkrimináció megelőzése, elszámoltathatóság biztosítása, felelős adatkezelés, az érdekelt hatásvizsgálata, méltányosság a végrehajtásban.
Ethics Guidelines For Trustworthy AI (2019)	egyéni autonómia, kármegelőzés, méltányosság, megmagyarázhatóság, átláthatóság, elszámoltathatóság	alapvető jogok, technikai megbízhatóság, adatkezelés, társadalmi hatás, környezeti jólét, az érdekelt felek bevonása
Recommendation on the Ethics of AI (2021)	emberi jogok, környezeti fenntarthatóság, inkluzivitás, átláthatóság, elszámoltathatóság, méltányosság, a magánélet védelme, etikai hatásvizsgálat	kulturális hatások, környezet- és ökoszisztéma-védelem, nemek közötti egyenlőség, oktatás és kutatás, adatkezelés, nemzetközi együttműködés
Initial policy considerations for generative AI (2023)	kockázatkezelés, társadalmi hatás, szellemi tulajdonjogok, dezinformáció-, előítéletesség- és diszkriminációellenesség, munkaerő-piaci hatások figyelembevétele	téves információk, szerzői jogi kérdések, elfogultságok csökkentése, munkahelyek automatizálása, túlzott támaszkodás a mesterséges intelligenciára, biztonsági aggályok
Hiroshima Process Guiding Principles (2023)	kockázatkezelés, átláthatóság, elszámoltathatóság, környezeti és fizikai biztonság, adatvédelem, felelősségteljes gazdálkodás, tartalomhitelesítés	AI életciklus-menedzsment, sebezhetőségi vizsgálatok, nemzetközi szabványok, társadalmi kockázatok enyhítése, globális kihívások kezelése, adatminőség, a szellemi tulajdon védelme
Recommendation of the Council on Artificial Intelligence (2024)	inkluzív növekedés, fenntartható fejlődés, jólét, emberi jogok, átláthatóság, megmagyarázhatóság, stabilitás, környezeti és fizikai biztonság, humán biztonság, elszámoltathatóság	kormányzás és politikai környezet, humán kapacitásépítés, munkaerőpiaci átalakulás, nemzetközi együttműködés, környezeti fenntarthatóság, kockázatkezelés, megbízható mesterséges intelligencia, érdekelt felek bevonása
EU AI Act (2024)	kockázat alapú megközelítés, emberközpontú mesterséges intelligencia, átláthatóság, biztonság, elszámoltathatóság, szabályok betartása	magas kockázatú mesterséges intelligenciát használó rendszerek, biometrikus azonosítás, adatkezelés, alapjogvédelem, innováció támogatása, fogyasztóvédelem, környezeti hatás

A táblázat elemzése nyomán számos jól kirajzolódó tendencia azonosítható a mesterséges intelligencia etikai kereteinek szabályozásában. Példaként említhetők az olyan kulcsfogalmak, mint az elszámoltathatóság vagy az emberi jogok tiszteletben tartása, amelyek visszatérő témák minden dokumentumban, az Asilomar-elvektől kezdve az EU AI Act-ig. Ezek jelzik, hogy

globálisan növekvő az egyetértés bizonyos elvek jelentőségét illetően, ahogy ezzel párhuzamosan megfigyelhető az átmenet az átfogó fókuszpontoktól a célzottabb területek felé. Például az Asilomar -elvek szélesebb körű etikai aggályokkal foglalkozott, míg az OECD és az EU későbbi iránymutatásai a hangsúlyt olyan speciálisabb területekre helyezték át, mint a munkaerőpiaci átalakulás, az adatvédelem és a demokratikus részvételre gyakorolt hatások szabályozása. Továbbá fontos megemlíteni, hogy egyre nagyobb hangsúlyt kapnak a környezeti fenntarthatósággal és a társadalmi hatásokkal kapcsolatos aggályok is, ami annak a felismerése, hogy az AI fejlődése nem kizárólag a technológiára, hanem a társadalomra és a környezetre is jelentős befolyással van. A dokumentumok kidolgozásában is észrevehető elmozdulást tapasztalhatunk egy komplexebb, az átfogóbb megközelítés felé, például az UNESCO ajánlásai vagy az EU AI Act már a mesterséges intelligencia teljes életútját felölelik, a tervezéstől és fejlesztéstől az alkalmazás hosszú távú következményeiig.

A továbbiakban az innováció és felelősség fogalmainak megjelenését szeretném megvizsgálni az adott dokumentumok kontextusában. Elsőként az innovációhoz való viszonyt vizsgáltam, amely során megállapítottam, hogy az elemzett dokumentumok egyértelműen kiállnak az innováció támogatása mellett, hangsúlyozva, hogy ez etikus keretek között, a társadalom érdekeit szem előtt tartva menjen végbe.

Az Asilomar-elvek (2017) ösztönzik a meghatározott etikai keretek között végbemenő innovációt, hangsúlyozva a közjó előmozdításának és a mesterséges intelligencia emberi értékekkel való összehangolásának fontosságát. Hasonlóképpen, az Understanding AI Ethics and Safety, amely egy olyan felelősségteljes innováció mellett érvel, ami garantálja a méltányosságot, az elszámoltathatóságot és a fenntarthatóságot. Az Ethics Guidelines For Trustworthy AI szintén hangsúlyozza, hogy az innovációnak összhangban kell lennie az etikai elvekkel, és az emberi autonómiára és a társadalmi jólétre kell összpontosítania. A Recommendation on the Ethics of AI üdvözli az innovációt az AI területén, a globális kihívások kezelésének eszközét látva benne, amennyiben biztosítja a fenntarthatóságot és a befogadó társadalom felépítését. A Initial policy considerations for generative AI kiemeli, hogy szükség van az innovációra, különösen a mesterséges intelligencia fejlesztése és alkalmazása során felmerülő kihívások kezelése terén. A The Hiroshima Process Guiding Principles a felelős innovációt helyezi fókuszba, amely révén a mesterséges intelligencia képes lesz hozzájárulni a globális kihívások megoldásához. A Recommendation of the Council on Artificial Intelligence úgyszintén támogatja az innovációt, feltéve, hogy az biztosítja a AI rendszerek megbízhatóságát, valamint az emberi jogokkal való összhangot. Végül az EU mesterséges

intelligenciáról szóló rendelte szintén szorgalmazza az innovációt, amennyiben az garantálja a közérdekek magas szintű védelmét, beleértve a környezeti és fizikai biztonságot és az alapvető jogokat.

7. táblázat. *Innováció és felelősség a dokumentumokban (saját szerkesztés)*

Dokumentum	Innováció	Felelősség
Asilomar Principles (2017)	Ösztönzi az innovációt az etikai határok figyelembevételével, hangsúlyozza a közös jólétet és a mesterséges intelligencia és az emberi értékek összehangolását.	Hangsúlyozza a felelősséget a fejlesztésében, kiemelve az etikai normákhoz és az emberi értékekhez való igazodást.
Understanding AI Ethics and Safety (2019)	Elősegíti a felelősségteljes innovációt, amely biztosítja a méltányosságot, az elszámoltathatóságot és a fenntarthatóságot.	Az AI rendszerek elszámoltathatóságát kiemelten kezeli, hangsúlyt helyez a felelős adatkezelésre és az etikus döntéshozatalra
Ethics Guidelines For Trustworthy AI (2019)	Az innovációnak összhangban kell lennie az etikai elvekkel, az emberi autonómiára és a társadalmi jólétre kell összpontosítania.	A felelősséget az elszámoltathatósághoz kapcsolja, továbbá a mesterséges intelligenciát használó rendszerek ne okozzanak kárt és átláthatóak legyenek.
Recommendation on the Ethics of AI (2021)	Ösztönzi az innovációt a mesterséges intelligencia területén a globális kihívások kezelése érdekében, amennyiben az biztosítja a fenntarthatóságot és a befogadást.	A felelősség központi szerepet játszik, hangsúlyozva, hogy a mesterséges intelligenciát használó rendszereknek elszámoltathatónak, átláthatónak és az emberi jogokkal összhangban állónak kell lenniük.
Initial policy considerations for generative AI (2023)	Kiemeli, hogy szükség van az innovációra a kockázatok kezelésében és a mesterséges intelligencia fejlesztése és alkalmazása során felmerülő kihívások kezelésében.	A felelősség magában foglalja a generatív mesterséges intelligenciával kapcsolatos kockázatok kezelését, beleértve a lehetséges társadalmi hatásokat és a visszaéléseket.
Hiroshima Process Guiding Principles (2023)	Támogatja a felelős innovációt a mesterséges intelligencia fejlesztésben, amennyiben hozzájárul a globális kihívások megoldásához	A felelősség annak biztosítását foglalja magában, hogy a mesterséges intelligenciát használó rendszerek biztonságosak, átláthatók és elszámoltathatók legyenek az életciklusuk során.
Recommendation of the Council on Artificial Intelligence (2024)	Támogatja az innovációt, amennyiben az biztosítja, hogy a mesterséges intelligenciát használó rendszerek biztonságosak, megbízhatók legyenek és összhangban álljanak az emberi jogokkal.	A felelősségvállalás fogalma az ellenőrzéshez kapcsolódik, és arra fókuszál, hogy az AI rendszereket etikusan és a szabályozásoknak megfelelően fejlesszék és használják.
EU AI Act (2024)	Támogatja az innovációt az AI területén, amennyiben az biztosítja a közérdekek magas szintű védelmét, beleértve a környezeti és fizikai biztonságot és az alapvető jogokat	A felelősséget a szabályozások betartása révén érvényesítik, a mesterséges intelligenciát használó rendszerek biztonságos és etikus használatára összpontosítva, különösen a magas kockázatú mesterséges intelligencia esetében.

A mesterséges intelligencia fejlesztésével és alkalmazásával kapcsolatos emberi felelősség kérdését az összes elemzett dokumentum hangsúlyozza. Az Asilomar-elvek kiemelik a kutatók

és fejlesztők felelősségét az etikus gyakorlatok betartásában a mesterséges intelligencia teljes életciklusa során. Az Alan Turing Institute által közreadott „Understanding AI Ethics and Safety” című dokumentum az AI-rendszerek kapcsán a felelősségre vonhatóság gondolatát emeli ki, és az emberi cselekvésre és a szándéokra helyezi a hangsúlyt. A fejlesztőkkel szemben elvárásként jelenik meg az, hogy garantálják az AI rendszerek működésének megfelelőségét a társadalmi értékrendnek, és nem ártás alapelveinek: *„It demands the active cooperation of all team members both in maintaining a deeply ingrained culture of responsibility and in executing a governance architecture that adopts ethically sound practices at every point in the innovation and implementation lifecycle”*³².(ATI, 2019, p. 5). Az „Ethics Guidelines for Trustworthy AI” is szorosan összekapcsolja a felelősséget az elszámoltathatósággal, hangsúlyozva, hogy az AI rendszereket úgy kell megtervezni és felhasználni, hogy elkerüljék a potenciális károk okozását. A fejlesztők és a szervezetek felelősséggel tartoznak e rendszerek biztonságának, jogszerűségének, etikusságának és megbízhatóságának biztosításáért: *„it is the responsibility of AI practitioners to ensure that they comply with their legal obligations, both as regards horizontally applicable rules as well as domain-specific regulation”*³³(EU, 2019, p. 15). Az UNESCO 2021-es „Recommendation on the Ethics of AI” című kiadványa a közös felelősségvállalás jelentőségét hangsúlyozza, globális párbeszédet szorgalmazva, mivel az AI etikai kihívásai nem kezelhetők elszigetelten. A szöveg kiemeli az emberi ágencia és felelősség kérdését: *„an AI system can never replace ultimate human responsibility and accountability”*³⁴(UNESCO, 2021, p. 13). Az ember csak korlátozottan adhat át döntéshozatali hatóköröket az AI-nak, a mesterséges intelligenciát használó rendszerek feletti kontrollt meg kell őrizni.

Az Initial Policy Considerations for Generative AI a félretájékoztató kapcsán hangsúlyozza, hogy nagy felelősség hárul a fejlesztőkre és az AI-t alkalmazó cégekre, amelyek sokszor az egyénekre hárítják a téves információk felhasználásából adódó felelősséget. Emiatt is figyelmet kell szentelni technológia társadalmi következményeinek és lehetséges visszaéléseinek szabályozására és olyan szakpolitikák kialakítására, amelyek érvényre juttatják a fejlesztők felelősségét az AI-rendszerek etikus használatáért. A Hiroshima Process

³² „Ez megköveteli a csapat minden tagjának aktív együttműködését, mind a mélyen gyökerező felelősségvállalási kultúra fenntartásában, mind pedig egy olyan irányítási struktúra megvalósításában, amely az innováció és a végrehajtás életciklusának minden egyes pontján etikailag megalapozott gyakorlatokat alkalmaz.” (saját fordítás)

³³ „a mesterséges intelligenciával foglalkozó szakemberek felelőssége annak biztosítása, hogy megfeleljenek jogi kötelezettségeiknek, mind a horizontálisan alkalmazandó szabályok, mind a terület-specifikus szabályozás tekintetében.” (saját fordítás)

³⁴ „egy mesterséges intelligencia rendszer soha nem helyettesítheti a végső emberi felelősséget és elszámoltathatóságot.” (saját fordítás)

International Guiding Principles egy átfogó, felelősségteljes megközelítést szorgalmaz az AI fejlesztésében és alkalmazásában, fókuszba állítva a mesterséges intelligenciát használó rendszerek biztonságának, átláthatóságának és elszámoltathatóságának garantálását a létrehozásuktól az életciklusuk lezárultáig. A dokumentumban ajánlás fogalmazódik meg a transzparenciáról, a fejlesztésben érintettek kockázatkezelési politikáinak nyilvánosságáról az elszámoltathatóság érdekében.

A Recommendation of the Council on Artificial Intelligence dokumentum kapcsolatban igyekszik teremteni az elszámoltathatóság és a kormányzás, valamint az etikai alapelveknek megfelelés között. Az OECD ajánlás kitér arra, hogy a kormányoknak felelősséget kell vállalniuk az AI-barát ökoszisztéma kialakításáért, a megfelelő szabályozási környezet megteremtéséért, valamint az AI generálta munkaerőpiaci változások kezeléséért – viszont a fejlesztők és fejlesztésben érintett cégek felelőssége alá tartozik a helytállás a mesterséges intelligencia jogi és etikai következményeiért.

EU mesterséges intelligenciáról szóló 2024-es rendelete hangsúlyozza, hogy a felelősség elve a mesterséges intelligencia szabályozásának egyik központi eleme, és felhívja a figyelmet a jogszabályok betartására, kiemelve a mesterséges intelligenciát használó rendszerek biztonságos és etikus használatát, különösen az úgynevezett magas kockázatúnak minősített rendszerek esetében. A rendelet részletesen szabályozza ezen rendszerek fejlesztőinek és forgalmazóinak felelősségét, azonban ezek felhasználóinak a felelőssége nem meghatározott, ahogy az egyéni felelősség kérdésköre is tisztázatlan marad.

5.1.4 Az etikai szabályozás hiányosságai

Elmondható, hogy a dokumentumok mindegyike kiemeli a felelősség jelentőségét a mesterséges intelligencia szabályozásával összefüggésben, és ezt a felelősséget gyakran az elszámoltathatóság kontextusában jelenítik meg. A felelősségvállalás mikéntje és gyakorlati oldala, sokkal kisebb hangsúlyt kap, leginkább az AI-rendszerek folyamatos ellenőrzése és értékelése iránti elkötelezettséget foglalja magában, annak biztosítása érdekében, hogy azok a megállapított szabályok keretein belül működjenek.

Amit a tárgyunk szempontjából fontos kiemelni, hogy bár ezek a dokumentumok alapos áttekintést – és egyes esetekben elemzést – adnak a mesterséges intelligencia által felvetett etikai dilemmákról, nem vesznek kellőképpen figyelembe felelősségtikai nézőpontokat. Az ajánlások és iránymutatások többsége elsősorban a közvetlen veszélyekre helyezi a hangsúlyt, és annak ellenére, hogy elvek szintjén hosszútávú folyamatok szabályozására tesznek kísérletet, kevesebb figyelmet fordítanak a mesterséges intelligencia potenciálisan a jövő generációkat

érintő, akár visszafordíthatatlan következményeire. A dokumentumokban jellemzően egy kötelességetikai megközelítés dominál, a legtöbb ajánlás olyan elveket fogalmaz meg, amelyeket *be kell tartani* (pl. átláthatóság, megkülönböztetésmentesség, elszámoltathatóság, munkaerőpiaci hatások figyelembevétele, a humán biztonság elve). Emellett hangsúlyosan jelen van egy következményetikai szemlélet, például több dokumentum is foglalkozik a mesterséges intelligencia potenciális negatív hatásaival és általánosságban az összes ajánlás kockázatközpontú szemlélettel bír. A dokumentumokban a közjó előmozdítására törekvés és a társadalmi haszon célként való megjelenítése, egy utilitarista szemléletet tükröz – ezek az ajánlások jellemzően azt szorgalmazzák, hogy a mesterséges intelligencia alkalmazásával, a felhasználók maximalizálják a társadalomra gyakorolt pozitív hatásokat, miközben minimalizálják a negatív következményeket.

A dokumentumokból egyértelműen hiányzik a felelősségetikai látásmód, egy hosszú távú hatások gondos mérlegelésén alapuló, a rendszerszintű átalakulásokkal gyakorlati megközelítésben számoló alapállás. A felelősségetika a technológia fejlődéséhez való proaktív és óvatos hozzáállás mellett érvel, vagyis ez a szemlélet nem csupán az AI közvetlen felhasználásának és hatásainak vizsgálatát foglalja magában, hanem egyben értelmezési keretet és programot is kínál. A felelősségetika arra hívja fel a figyelmet, hogy minden innovációs folyamat – így az AI fejlesztések előtt is – mérlegelnünk kell a hatások komplex mátrixát. Milyen lehetséges környezeti következményei lehetnek az AI-rendszerek növekvő energiaigényének? Milyen hatással lehet a mesterséges intelligencia az emberi kultúra és az értékek fejlődésére, mennyire képes átformálni az emberiség globális kulturális arculatát? Ezt követően kell cselekvési terveket kidolgozni, amely kereteinek meghatározást az elemzésünk eredményei segítik.

Az említett hiányosságokra példaként – esettanulmányként – érdemes megvizsgálni az Európai Parlament és a Tanács 2024/1689 rendeletét, mint a mesterséges intelligencia a világon egyedülálló szabályozását. A részletes elemzés során jól megállapíthatók a fent említett hiányosságok: a rendelet nem foglalkozik kellő mélységében a hosszú távú következményekkel – habár szellemiségében implicit módon jövőorientált – de szövegében kifejezetten a rövid távú kockázatokra, a jelenlegi piaci és technológiai helyzetre koncentrált, példa erre a bevezetés első szakasza: „*E rendelet célja, hogy javítsa a belső piac működését, valamint, hogy előmozdítsa az emberközpontú és megbízható mesterséges intelligencia (MI) elterjedését*” (EU AI Act, 2024, p. 1). A szöveg egészében látenciában marad a jövőbeli következmények figyelembevétele, a rendelet összes szakasza a jelenben folyó kutatási és innovációs, illetve

kereskedelmi tevékenységre fókuszál. A rendelet nem tartalmaz olyan szabályokat, amelyek az AI-rendszerek fejlesztését megelőzően előírná a lehetséges káros következmények azonosítását és számbavételét, és azok mérséklésére alkalmazandó eljárásokat és tevékenységeket. A szöveg egészében nem jelenik meg olyan követelmény, amely arra kötelezné a fejlesztőket és a felhasználókat, hogy proaktívan felmérjék a potenciális negatív hatásokat és azokat a forgalomba hozatal előtt kiküszöböljék. Ahogy az etikai elvek is általánosságok szintjén fogalmazódnak meg: „biztosítja az etikai elvek védelmét” (p. 2), „bármely kutatás-fejlesztési tevékenységet a tudományos kutatásra vonatkozó elismert etikai és szakmai normáknak megfelelően kell végezni” (p. 7) továbbá (a szövegben utalva a 2019-es fent is tárgyalt iránymutatásokra) leszögezi, hogy az „említett elvek alkalmazását **lehetőség szerint át kell ültetni az MI-modellek tervezésébe és használatába**” és „**valamennyi érdekelt fél – ideértve az ipart, a tudományos köröket, a civil társadalmat és a szabványügyi szervezeteket is – ösztönözve van arra, hogy az önkéntes legjobb gyakorlatok és szabványok kidolgozásához adott esetben vegye figyelembe az etikai elveket**” (p. 8). A kiemelések rámutatnak a szabályozásban szereplő elvek alkalmazásának opcionális kezelésére, amit további szakaszok is megerősítenek, például a rendelet tizedik „Magatartási kódexek és iránymutatások” fejezetének, 95. „A konkrét követelmények önkéntes alkalmazására vonatkozó magatartási kódexek” cikkének második szakasza: „Az MI-hivatal és a tagállamok elősegítik olyan magatartási kódexek kidolgozását, amelyek a konkrét követelmények valamennyi MI-rendszerre való – többek között az alkalmazók általi – **önkéntes alkalmazására vonatkoznak**”. Az utalás a magatartási kódexekre több helyen előfordul a rendelet szövegében (pl. a 20., 27., 165. szakasz), azonban konkrét tartalmuk meghatározatlan marad, míg a 96. cikk 4. szakaszában megfogalmazottak alapján kiderül, hogy ezek tartalmi elemeinek meghatározásánál sincsenek gyakorlati alkalmazásba átültethető alapelvek: „A magatartási kódexek kidolgozásának **ösztönzése és elősegítése során az MI-hivatal és a tagállamok figyelembe veszik a kkv-k – köztük az induló innovatív vállalkozások – sajátos érdekeit és igényeit**”.

A szöveg több helyen említ szintén közelebről nem definiált teljesítménymutatókat, amelyek kapcsán korábban azonosított alapelveket (AI-ismeretek, exkluzivitás és megkülönböztetésmentesség, nemi egyenlőség, akadálymentesség stb.) említ, kiemelve olyanokat, mint a „**az MI-rendszerek környezeti fenntarthatóságra gyakorolt hatásának értékelése és minimalizálása**” vagy „**azon negatív hatások értékelése és megelőzése, amelyeket az MI-rendszerek a kiszolgáltató személyekre vagy a kiszolgáltató személyek csoportjaira (...) valamint a nemek közötti egyenlőségre gyakorolnak**” (EU AI Act, 2024, p. 113).

Elmondható, hogy a rendelet elsősorban a mesterséges intelligencia szabályozásának technikai és eljárási aspektusaira koncentrált, illetve ezeknek való megfelelésre és teljes mértékben adós marad az etikai kérdések vizsgálatával. A szövegben sem a globális hatások mérlegelése, sem az elővigyázatosság elve nem fogalmazódik meg explicit módon, az etikai alapelvek utalás szintjén kerülnek felsorolásra azok figyelembevételét opcionálisan kezelve, a konkrét magatartási kódexek kidolgozásának kérdését a jövőbe utalva.

Határozott véleményem, hogy a felelősségetikai szempontok beépítése a mesterséges intelligencia etikai szabályozásának keretébe, előremutató és hatékony ellenőrzési rendszer létrehozására ad lehetőséget. Mielőtt a felelős döntéshozatal-támogató elemzési keretet bemutatnám, röviden áttekintem azokat a fő területeket, amelyek a legjelentősebb etikai kihívásokkal rendelkeznek. Ezek meghatározásában segítségemre voltak az eddigiekben tárgyalt dokumentumok, amelyek alapján négy problémakört határoztam meg: [1] biztonságpolitikai kihívások, [2] globális adathálózatok által felvetett problémák, [3] a munkaerőpiacot érintő változások, és [4] az oktatási rendszer átalakulása.

5.2 Etikai kihívások a mesterséges intelligencia alkalmazásának területén

5.2.1 Az innováció kihívásai a biztonságpolitika területén

Az előző fejezetben kitértem a felelős innováció kérdéskörére, amely elv alapján a kutató a fejlesztés során figyelembe veszi a társadalmi, környezeti és etikai hatásokat és kockázatokat, beépítve azokat a szemléletmódokat, amelyek a fejlesztések lehetséges jövőbeli következményeit széleskörűen elemezni képesek. Biztonságpolitikai megközelítésből a felelős innováció az olyan technológiai fejlesztéseket és alkalmazásokat öleli fel, amelyek egyszerre veszik figyelembe a társadalmi, etikai és biztonsági kérdéseket. Ez a megközelítés biztosítja, hogy az új technológiák nem csak hatékonyak és előremutatók, hanem erkölcsileg is elfogadhatók és társadalmilag felelősek is lesznek egyúttal.

A technológiai fejlődés felgyorsulásával, az innováció és a biztonság (legszélesebb értelmében vett jelentése) közötti bonyolult dinamika megértése napjainkra kulcsfontosságúvá vált, a fejlődés, valamint a stabilitás fenntartása és a fenyegetések elhárítása közötti egyensúly fenntartása szempontjából. A probléma lényege az innováció gyors üteme és a biztonságpolitika kialakításának hagyományosan mérlegelő jellege közötti feszültségben rejlik. A biztonságpolitikai tervezés és stratégiák gyors technológiai fejlődés nyomán történő késedelmes kiigazítása olyan potenciális sebezhetőségekhez vezethet, amelyeket ellenséges hatalmi tömbök, ellenérdekelt szereplők kihasználhatnak. E probléma jelentősége túlmutat a

közvetlen biztonsági következményeken, és tágabb geopolitikai, társadalmi-gazdasági és etikai dimenziókat is magában foglal.

Ahogy a technológiai innováció rohamléptekkel halad előre, egymás után jelennek meg olyan etikai természetű kérdések, amelyekkel már a problémák korai szakaszában érdemes a döntéshozóknak foglalkozniuk. Bár a fejlesztések önmagukban értéksemlegesek, alkalmazásuk a komplex társadalmi keretek között, sokrétű etikai megfontolásokat vet fel.

Az egyik ilyen kérdés a mesterséges intelligenciához kapcsolódóan a döntési felelősség és szabadság kérdése. A mesterséges intelligencia egészségügyi diagnosztikától az igazságszolgáltatási rendszerekig terjedő, különböző ágazatokban megfigyelhető térhódítása óhatatlanul felvet kérdéseket a döntéshozatali szabadság algoritmusoknak való átengedésének erkölcsi következményeiről. Kérdések merülnek fel az elszámoltathatósággal, az átláthatósággal és bizonyos előítéletek lehetséges állandósulásával kapcsolatban (Boddington, 2017). Az emberi életet befolyásoló döntések mesterséges intelligenciának történő delegálása – legyen szó akár az egészségügyről vagy igazságszolgáltatásról – már önmagában is megkérdőjelezi az emberi méltóság, a cselekvőképesség és a személyi szabadság alapelveit (Mittelstadt, 2019).

Hasonló erkölcsi problémákkal szembesülünk az IoT-eszközök (okosórák, biztonsági kamerák, világítóeszközök, közlekedési szenzorok, egészségügyi monitorok stb.) elterjedésével kapcsolatban, amelyek miközben hatékonyságot, gyors hozzáférést és számos kényelmi szolgáltatást ígérnek, egyúttal a magánélet határait elmoszák. A személyes adatok gyűjtése, tárolása és elemzése nemcsak az egyént teszi ki a potenciális visszaélésnek, hanem a személyes terek, élmények, privát kapcsolatok minőségét is erodálja, kihívást jelentve a normatív adatvédelmi etikának (Zuboff, 2019). A tárgyak internetének egyre növekvő megfigyelési potenciálja tovább fokozza az ilyen adatok manipulatív vagy zsarolási célokra való felhasználásával kapcsolatos aggodalmakat (Schneier, 2015).

A digitális pénzügyi tranzakciók nyilvántartására és biztonságos kezelésére kifejlesztett blokklánc technológia egyik legfőbb erősségének tartják azt, hogy az adatok nem módosíthatók, lényegében megváltoztathatatlanok. Miközben biztosítja az adatok integritását, ez azt is jelenti, hogy a potenciálisan hibás vagy káros rögzített információk, ha egyszer hozzáadódnak, visszavonhatatlanok. Olyan kontextusokban, ahol a javítás vagy törlés etikailag vagy jogilag indokolt lehet, ez jelentős etikai kihívásokat jelent, és az ilyen helyzetek kezelésével számolni kell a biztonsági szakembereknek és a törvényalkotóknak egyaránt (Maurer et al., 2013). Ugyanígy, például a jövőben, ha a kvantumszámítógépek képesek lesznek feltörni a mai

titkosításokat, az eddig védett információk sokasága kerülhet napvilágra, amely szintén komoly etikai kérdéseket vet fel az információbiztonság és a magánélet védelmével kapcsolatban. (Aaronson, 2013).

Az egyéni jogok és a közösségi biztonság közötti feszültség egyidős a kormányzás modern fogalmával. Míg a közösség védelmének imperatívusza aláhúzza bizonyos korlátozások indokoltságát, e korlátozások mértéke és az egyéni jogokkal való ütközésük továbbra is szigorú etikai kérdéseket vet fel. Ahogy a következő fejezetekben tárgyalni fogom részletesen, az információs korban folyamatosan hatalmas mennyiségű személyes adatot generálunk, gyűjtünk és cserélünk. Az egyének egyre több személyes információt adnak át állami és vállalati szervezeteknek, az adatok kezelésének, a hozzáférésnek és az esetleges felhasználásnak az etikai dimenziói kiemelkedő jelentőségűvé válnak. Az, hogy miként, miért és ki használja fel ezeket az adatokat, növelheti vagy alááshatja az egyéni jogokat és a kollektív biztonságot (Zuboff, 2019). A tömeges adatgyűjtés, függetlenül a konkrét fenyegetésektől, komoly aggodalmakat vet fel a magánéletbe való indokolatlan behatolással kapcsolatban (Richards, 2013).

Az olyan technológiai fejlesztések, mint a mesterséges intelligencia által vezérelt arcfelismerő rendszerek, bár ellenőrzött biztonsági protokollokkal dolgoznak, de ennek ellenére a téves azonosítás, a kódolt előítéletek és az indokolatlan megfigyelés kockázatát továbbra is magukban hordozzák. E technológiák folyamatos jelenléte a nyilvános terekben megkérdőjelezi a magánülethez, a privát szférához való jogokat, mivel az egyének folyamatosan egy algoritmus vezérelte ellenőrzésnek vannak kitéve.

Az összekapcsolt világban a technológiai fejlődés hatásai átlépik az országhatárokat. Az innováció dinamizmusa, számtalan ígéretével és buktatójával, szorosan kapcsolódik a globális biztonság változó helyzetéhez. Mivel az államok egyre inkább a technológiára támaszkodnak a kormányzás, a védelem és az állampolgári folyamatok terén, bármilyen meghibásodás földomlásszerű katasztrófákhoz és hosszútávú krízisekhez vezethet. Az IoT technológiák globális térhódítása, bár innovatív kényelmi megoldásokat kínálnak, ugyanakkor felerősíthetik a rendszerszintű sebezhetőségeket, ha nem biztosítják őket megfelelően (Owens et al., 2019). A mesterséges intelligencia által vezérelt autonóm fegyverrendszerek megjelenése széles körű vitákat vált ki etikai, jogi és biztonsági szempontból. Miközben növelhetik a pontosságot és hatékonyságot, csökkentve az emberi áldozatok számát, aggodalmakat is felvetnek egy esetleges nem szándékos eszkaláció, az elszámoltathatóság hiányosságai és a nem állami szereplők ezen fegyverrendszerekhez való hozzáférése miatt (Scharre, 2018). A digitális

platformok, annak ellenére, hogy demokratizálják az információterjesztést, fegyverként szolgálnak a dezinformációk terjesztésére, a vélemények manipulálására és a demokratikus intézmények aláásására. A *fake news* kampányok globális hulláma megerősítette ezt a kihívást (Wardle & Derakhshan, 2017).

Az elmúlt évtizedek technológiai innovációi transzformatív jellegűek, azaz a kultúrában, a társadalomban és a gazdaságban is mélyreható változásokat idéznek elő. A mesterséges intelligenciától a kvantumszámítógépekig, ezek az áttörések olyan módon alakítják át a világunkat, amelyet még csak most kezdünk felfogni. Miközben azonban ezeket a technológiákat integráljuk társadalmi rendszereinkbe, meg kell küzdenünk a bevezetésükkel megjelenő veszélyekkel és kihívásokkal is. A kiberbiztonsági aggályok, az adatvédelem megsértése, sőt a fizikai biztonságot fenyegető veszélyek is felmerülnek, ahogy a technológiailag lehetséges határait feszegetjük (Smith & Brooks, 2015).

A biztonságpolitikát érintő döntéshozatali folyamatokban várhatóan kiemelt szerepet kap a már említett felelős innováció kérdése. A felelős innováció olyan inkluzív tervezést jelent, amely biztosítja, hogy a technológia a társadalom minden szegmensét szolgálja, és véletlenül se marginalizáljon vagy hozzon hátrányos helyzetbe egyetlen csoportot sem (Dignum, 2019). A technológiai kihívások határok nélküli jellege miatt az etikai megoldások nemzetközi együttműködést, közös normákat és szerződéseket tesznek szükségessé, amelyek biztosítják a kölcsönös biztonságot és az emberi jogok védelmét. Az összekapcsolt világban a biztonsági kihívások és a technológiai újítások gyakran átlépik a határokat. A közös etikai iránymutatások és nemzetközi normák kidolgozása kulcsfontosságúvá válik annak biztosításában, hogy az innovációkat a kölcsönös előnyök és nem a kölcsönös pusztítás érdekében használják fel. Ahogy az államok fontolóra veszik a mesterséges intelligencia által vezérelt rendszerek felhasználását a védelemben, etikai kérdések merülnek fel azzal kapcsolatban, hogy milyen erkölcsi következményekkel járhat, ha élet-halál kérdésekben döntéseket bízunk a gépekre. Az emberi felügyelet és az erkölcsi felelősség biztosítása kiemelkedően fontossá válik (Russell et al., 2015). A fejlett algoritmusok és a gépi tanulási technikák felismerhetik a fenyegetéseket, megjósolhatják a bűncselekményeket, vagy segíthetnek az összetett biztonsági döntések meghozatalában. E technológiák jellege azonban olyan etikai megfontolásokat igényel, amelyek elősegítik az átláthatóságot, biztosítva a folyamatok érthetőségét és elszámoltathatóságát (Bostrom & Yudkowsky, 2014).

Az innováció és a biztonságpolitika dinamikus kölcsönhatásában hatalmas lehetőségek és példátlan kihívások kereszteződésében találjuk magunkat. Az innováció a maga sokrétű

dimenzióiban új lehetőségeket kínál biztonsági paradigmáink javítására és újradefiniálására, de legalább ugyanennyi kérdést is felvet: a haladás összetett etikai dilemmákkal és a bevett biztonsági normák potenciális veszélyeztetésével jár együtt.

5.2.1.1 A szingularitás és az AI fejlődésének dilemmái

A mesterséges intelligencia elmúlt években tapasztalható robbanásszerű fejlődése számos spekulációt eredményezett a közeli jövőben bekövetkező *technológiai szingularitás* lehetőségéről – vagyis egy olyan feltételezett időpontról, amikor az AI meghaladja az emberi kognitív kapacitásokat és képessé válik önmagát módosítani és fejleszteni, ami a jelenlegi felfogóképességünket meghaladó robbanásszerű technológiai fejlődéshez vezet el (Bostrom, 2014; Kurzweil, 2005). A szingularitás bekövetkezése egy mérföldkövet jelentene a technológia fejlődésében, mivel a gépeket többé nem korlátoznák az emberi intelligencia határai. Ahogy a mesterséges intelligencia egyre fejlettebbé válik, és képes lesz önmagát tovább tervezni és tökéletesíteni, ez a „rekurzív önfejlesztésnek” (Recursive Self-Improvement – RSI) nevezett folyamat olyan mesterséges intelligencia rendszereket eredményezhet, amelyek nemcsak intelligensebbek az embernél, hanem képesek olyan ütemben fejlődni és alkalmazkodni is, amely folyamat bekövetkeztére jelenleg nincsenek meg sem a technológiai, sem az etikai válaszaink (Yudkowsky, 2013).

Az általános mesterséges intelligencia (AGI – Artificial General Intelligence) a mesterséges intelligencia egy olyan formája, amely funkcionálisan hasonló az emberi intelligenciához, széleskörű feladatokat képes ellátni és komplex problémákat megoldani (Bubeck et al., 2023). Szemben a legtöbb jelenlegi mesterséges intelligenciát használó megoldással, amelyek specifikus feladatok megoldására korlátozódnak, az AGI olyan képességekkel rendelkezik, ami lehetővé teszi, hogy tanuljon, alkalmazkodjon és alkotó módon gondolkodjon és ezáltal minden olyan szellemi munkát el tudjon végezni amire az emberek képesek (Voss, 2007). Több cég dolgozik azon, hogy létrehozzon egy ilyen modellt, és a ChatGPT modellt kifejlesztő OpenAI egyes iparági információk alapján (Efrati & Victor, 2023) már 2023 nyarán jelentős áttörést ért el egy szuperintelligencia kifejlesztése terén.

Nick Bostrom a *Superintelligence: Paths, Dangers, and Strategies* (Bostrom, 2014) című írása kimerítően vizsgálja a mesterséges intelligencia (AI) lehetséges jövőjét a szuperintelligencia fogalmára összpontosítva, amely egy olyan mesterséges intelligencia, amely szinte minden gazdaságilag értékes feladatban felülmúlja az emberi intelligenciát. Bostrom állítja, hogy egy szuperintelligens AI létrehozása lehetséges és valószínű is a jövőben, és ha egyszer ezt kifejlesztettük, az még hatékonyabb gépek tervezésére lesz alkalmas, ahogy

ezt a matematikus I. J. Good –aki elsőként felismerte a szuperintelligencia kapcsán jelentkező veszélyeket – egy korábbi tanulmányában már leírta:

„an ultraintelligent machine could design even better machines; there would then unquestionably be an “intelligence explosion,” and the intelligence of man would be left far behind (see for example refs. Thus the first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control”³⁵ (Good, 1965, p. 33).

Bostrom számára az egyik központi kérdést a szuperintelligens AI-al kapcsolatos lehetséges kockázatok jelentik, véleménye szerint ez, ha nem megfelelően ellenőrzik és vezetik, egzisztenciális fenyegetést jelenthet az emberiségre nézve. Ennek oka, hogy az AI céljai ellentétbe kerülhetnek az emberi elvekkel és érdekekkel, és a magasabb intelligenciaszintjét felhasználhatja céljai elérésére – az emberi jólét rovására (Szilágyi, 2024b, p. 209). A kockázatok csökkentésére Bostrom számos megoldást javasol, többek között a szuperintelligens AI emberi felügyelet alatt tartását biztosító ellenőrzési módszerek kidolgozását, valamint az AI céljainak az emberi normákhoz és értékekhez való igazítását, vagyis elengedhetetlenek az olyan szabályozások, együttműködési protokollok és etikai keretrendszerek kialakítása, amelyek biztosítják, hogy a technológiák az ember javát szolgálják. Hangsúlyozza továbbá a globális együttműködés jelentőségét a mesterséges intelligencia fejlesztésének irányításában és kiemeli a gépi intelligencia kapcsán a biztonsági kockázatok kezelésének fontosságát, hangsúlyozva, hogy a kutatóknak el kell kötelezniük magukat a fejlesztések és alkalmazás során is a humán biztonság prioritása mellett (Bostrom, 2014, p. 358).

A szuperintelligencia problematikája rávilágít, hogy a technológiai fejlődésben rejlő lehetőségek mellett számot kell vetnünk a kockázatokkal is, és olyan megközelítést kell találnunk, mely egyensúlyt teremt az innováció és az emberi értékek védelme között. Csak így kerülhetjük el a szuperintelligencia potenciálisan katasztrofális hatásait. Erre mutat rá David Chalmers is, hogy az AI fejlesztése során az etikai megfontolásoknak központi szerepet kell játszaniuk, és olyan rendszereket kell létrehozunk, amelyek összhangban vannak az emberi értékekkel. Ez azonban számos filozófiai kérdést vet fel az AI tudatosságával, morális státuszával és az emberi ágens megőrzésével kapcsolatban. Izgalmas gondolat, amit az emberi agyat emuláló rendszerek (human-based AI) kapcsán felvet, amelyekben az értékek várhatóan

³⁵ „(...) Egy ultraintelligens gép még jobb gépeket tervezhetne; és ekkor kétségtelenül bekövetkezne egy „intelligencia-robbanás”, és az ember intelligenciája messze lemaradna. Így az első ultraintelligens gép az utolsó találmány, amit az ember valaha készítene, kivéve, ha a gép eléggé engedelmes ahhoz, hogy elmondja nekünk, miként tartsuk azt irányításunk alatt.” (saját fordítás)

hasonlóak lesznek az emberi értékekhez. A nem-emberi AI (non-human AI) esetében viszont gondoskodni kell a megfelelő értékek beépítéséről – itt az értékek pontos meghatározása és specifikálása nagy körültekintést igényel a későbbi nem kívánt következmények elkerülése érdekében – figyelmeztet (Chalmers, 2006, pp. 24-29).

5.2.2 A globális adathálózatok etikai kihívásai

A mesterséges intelligencia és általánosságban a technológia gyors fejlődése és mindenhol tapasztalható jelenléte olyan újszerű gazdasági logika kialakulásához vezetett el, amely messzemenő következményekkel jár az egyéni autonómiára, a magánéletre és társadalmunk szerkezetére nézve egyaránt.

Shoshana Zuboff a *The Age of Surveillance Capitalism* című művében úgy határozza meg ezt az új gazdasági rendet, amely, „*claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales; (...) A parasitic economic logic in which the production of goods and services is subordinated to a new global architecture of behavioral modification*”³⁶ (Zuboff, 2019, p. 9). Azt állítja, hogy a felügyeleti kapitalizmus mélyreható változást jelent a piac és az egyének közötti viszonyban, amelyben az emberi tapasztalat árucikké válik és pénzzé teszik, érdemi beleegyezésünk vagy kölcsönös előnyök nélkül. Zuboff tézisének középpontjában a prediktív analízisen alapuló értékesítés fogalma áll: ez a gépi tanulási algoritmusok használatára utal, amelyek hatalmas mennyiségű viselkedési adatot elemeznek a jövőbeli fogyasztói viselkedés előrejelzése és befolyásolása érdekében. Ahogy Zuboff megjegyzi, hogy a felügyeleti kapitalizmus felfedezte miként avatkozzon be az emberi viselkedésbe, miként formálja azt, az embereket hogyan terelje és hangolja a számára előnyös viselkedésmódok irányába (Zuboff, 2019, p. 14). Ezekben a modelleken alapuló példátlan lehetőségek az emberi viselkedés befolyásolására és fogyasztási célú alakítására mély etikai kérdéseket vetnek fel a szabad akaratról, az autonómia természetéről és a technológia az életünkben betöltött szerepéről.

A felügyeleti kapitalizmus középpontjában a felhasználók online tevékenységeiből származó adatok kinyerése áll. Ezeket az adatokat különböző technikákkal és módszerekkel gyűjtik, például nyomkövető cookie-kkal, mobilalkalmazásokkal és közösségi médiaplatformokon keresztül, amelyek lehetővé teszik a vállalatok számára, a felhasználók online viselkedésének, preferenciáinak és interakcióinak nyomonkövetését (Fuchs, 2014).

³⁶ „az emberi tapasztalatot ingyenes nyersanyagként tekinti a kitermelés, előrejelzés és értékesítés rejtett kereskedelmi gyakorlataihoz; (...) egy parazita gazdasági logika, amelyben az áruk és szolgáltatások előállítását a viselkedésmódosítás új globális architektúrájának van alárendelve” (saját fordítás)

Különösen a közösségi médiaplatformok váltak az emberi viselkedés alapján megszerzhető többletinformációk kinyerésének kulcsfontosságú helyszíneivé: „*social media use capital accumulation models that are based on the exploitation of the unpaid labour of Internet users and on the commodification of user-generated data and data about user behaviour that is sold as commodity to advertisers*”³⁷ (Fuchs, 2014, p. 122). Azzal, hogy látszólag ingyenes szolgáltatásokat kínálnak a felhasználói viselkedés megfigyelésének, elemzésének és pénzértételének jogáért cserébe, a közösségi média vállalatok a munkaerő kizsákmányolásának új formáját hozták létre, amely az átlagfelhasználó számára gyakran láthatatlan. Ráadásul a közösségi médiaplatformok kialakítása maga is a felhasználói elkötelezettség és az adatszerzés maximalizálására irányul, gyakran az egyéni jólét és a társadalmi kohézió rovására. A közösségi média függőséget okozó jellege, valamint az online élményeinket áruként kezelő átláthatatlan algoritmusok komoly kérdéseket vetnek fel azzal kapcsolatban, hogy gondolatainkat, érzéseinket és viselkedésünket milyen mértékben befolyásolják és manipulálják a kereskedelmi érdekek. Az összegyűjtött adatokat fejlett algoritmusok és gépi tanulási modellek segítségével dolgozzák fel és elemzik, hogy piacosítható előrejelzéseket készítsenek a felhasználói viselkedésről. Ahogy Yeung (2017) megjegyzi

*„(...) that the analytic phenomenon currently known as ‘Big Data’ can be understood as a mode of ‘design-based’ regulation (...) By highlighting correlations between data items that would not otherwise be observable, these techniques are being used to shape the informational choice context in which individual decision-making occurs, with the aim of channeling attention and decision-making in directions preferred by the ‘choice architect’”*³⁸ (Yeung, 2017, p. 118).

Ezeket az információkat, megfigyeléseket, előrejelzéseket aztán célzott reklámok, személyre szabott tartalomkínálat és más, a fogyasztói döntések és viselkedés befolyásolására irányuló stratégiák megalkotására használják.

Az egyéni viselkedés megfigyeléséből kapott adatokból nyert ismeretek felhasználásával a cégek irányíthatják a felhasználókat az általuk kívánt irányba: például vásárlásra ösztönözhetik az embereket hirdetési felületek megnyitására, vagy arra, hogy több

³⁷ „a közösségi média olyan tőkefelhalmozási modelleket alkalmaz, amelyek az internetfelhasználók fizetetlen munkájának kihasználásán, valamint a felhasználók által generált adatok és a felhasználói viselkedésre vonatkozó adatok árucikké tételén alapulnak, amelyeket árucikként adnak el a hirdetőknak.” (saját fordítás)

³⁸ „a jelenleg Big Data néven ismert jelenséget a ‘tervezésen alapuló’ szabályozás módjaként lehet értelmezni. Azáltal, hogy kiemelik az olyan adatelemek közötti összefüggéseket, amelyek egyébként nem lennének megfigyelhetők, ezeket a technikákat arra használják, hogy átalakítsák azt a választási kontextust, amelyben az egyéni döntéshozatal történik, azzal a céllal, hogy a figyelmet és a döntéshozatalt az általuk preferált irányokba terelessék.” (saját fordítás)

időt töltsenek egy platformon. A fogyasztói magatartás ilyen szintű manipulálása komoly etikai aggályokat vet fel az egyéni autonómia csorbulásával és a kizsákmányolás lehetőségével kapcsolatban. Ráadásul a felügyeleti kapitalizmusban használt algoritmusok és modellek áttekinthetlensége és megismerhetetlensége megnehezíti az egyének számára, hogy megértsék, hogyan használják fel az adataikat, és hogyan befolyásolják a viselkedésüket. Pasquale *black box* elmélete szerint a modern gépi tanulási rendszerek, egyfajta „fekete dobozként” működnek. A bemeneti adatok (*input*) és a kimeneti eredmény (*output*) ismert, de hogy mi történik a kettő között, az rejtve marad. Mivel nem tudjuk pontosan, hogyan és miért hoz meg egy AI rendszer bizonyos döntéseket, ezért nehéz az elszámoltathatóság. Az elmélet felhívja a figyelmet a modern AI megoldások átláthatatlanságára, és szorgalmazza, hogy kezeljük kiemelt ügyként az elszámoltatható, transzparens mesterséges intelligencia fejlesztését. Ahogy Pasquale érvel: „*The point of black boxes is to hide critical facts about what is going on. They undermine any confident assertions about the precise nature of firms' investment, accounting, or documentation practices*”³⁹ (Pasquale, 2015, p. 137). Az átláthatóság és az elszámoltathatóság hiánya tovább súlyosbítja a felügyeleti kapitalizmus által felvetett etikai dilemmákat.

5.2.2.1 A felügyeleti kapitalizmus társadalmi és demokratikus következményei

Jelentős problémaként azonosíthatók az online pszeudo-nyilvános terek, az olyan digitális platformok, amelyek nyitottnak és a nyilvánosság számára hozzáférhetőnek tűnnek, de valójában magántulajdonban vannak és cégek irányítják őket. Ezek a terek, mint például a már említett közösségi médiafelületek, keresőmotorok és digitális piacterek, döntő szerepet játszanak a viselkedési adatok kinyerésében és a felhasználói viselkedés manipulálásában. A pszeudo-nyilvános terek ezeket az aszimmetriákat példázzák azzal, hogy semlegesnek és felhasználó-központúnak mutatják magukat, miközben egyidejűleg kiterjedt adatgyűjtést és tartalomszabályozási tevékenységet folytatnak. A felhasználók gyakran nincsenek tisztában azzal, hogy milyen mértékben gyűjtik az adataikat, és hogyan használják fel azokat viselkedésük befolyásolására. Fuchs a közösségi médiáról írt könyvében a felhasználók „kettős árucikké válását” tárgyalja: „*they are commodities themselves and through this commodification their consciousness becomes, while online, permanently exposed to*

³⁹ „*A fekete dobozok lényege, hogy elrejtsek a kritikus tényeket arról, hogy mi történik. Aláássák a cégek befektetési, számviteli vagy dokumentációs gyakorlatának pontos természetére vonatkozóan kijelenthető állításokat.*” (saját fordítás)

*commodity logic in the form of advertisements. Most online time is advertising time*⁴⁰ (Fuchs, 2014, p. 169). Az ezeken a platformokon a felhasználói interakciók által generált adatokat a platformokat birtokló vállalatok gyűjtik, elemzik és monetizálják. E logikában a Google-hoz hasonló keresőmotorok és az Amazonhoz hasonló digitális piacterek is pseudo-nyilvános térként működnek: látszólag ingyenes szolgáltatásokat és termékeket kínálnak, miközben hatalmas mennyiségű adatot gyűjtenek a felhasználók kereséseiről, böngészési előzményeiről és vásárlási szokásairól. A felhasználók által látható információk és a felkínált lehetőségek ellenőrzésével ezek a platformok célzottan befolyásolhatják a felhasználói viselkedést és döntéshozatalt.

A fentiekben vázolt globális rendszer jelentős kihívást jelent az egyéni autonómia számára, mivel az adatgyűjtés és manipuláció átláthatatlan folyamatain keresztül működik, amelyek korlátozzák a felhasználók képességét arra, hogy tájékozott és megalapozott döntéseket hozzanak. Az átláthatóság hiánya, a döntéshozatalra gyakorolt befolyás és a magánélethez fűződő jogok csorbulása mind hozzájárulnak a személyes cselekvőképesség csökkenéséhez. Zuboff megjegyzi: „*Surveillance capitalists know everything about us, whereas their operations are designed to be unknowable to us*”⁴¹ (Zuboff, 2019, p. 16), a felhasználók gyakran nincsenek tisztában azzal, hogy személyes adataikat hogyan kezelik, és kevéssé tudják ellenőrizni, hogy ezeket az információkat hogyan használják fel online élményeik alakítására és viselkedésük befolyásolására. A felhasználóknak kínált információk és lehetőségek célzott kezelésével a platformok manipulálhatják a döntéshozatali folyamatokat, és olyan döntések felé terelhetik az egyéneket, amelyek inkább a vállalatok érdekeit szolgálják, mint a felhasználókét.

A felügyeleti kapitalizmus tényerése messzemenő következményekkel jár a társadalomra és a demokráciára nézve is, mivel a gazdasági és politikai hatalmat néhány domináns technológiai vállalat kezében összpontosítja, aláássa a szabad versenyt és a demokratikus felügyeletet, és egyenlőtlen hatalmi dinamikát teremt ezen óriások és a szolgáltatásaikra támaszkodó felhasználók között. Az egyik legjelentősebb társadalmi és demokratikus következmény a szabad verseny és a demokratikus felügyelet aláásása. Az olyan vállalatok, mint a Google, a Facebook és az Amazon piaci dominanciája lehetővé teszi számukra, hogy a digitális világot, a globális digitális környezetet a saját érdekeiket szolgáló módon alakítsák, gyakran a felhasználók magánéletének, autonómiájának és választási

⁴⁰ „Amelyekben a felhasználók maguk is árucikkek, és tudatuk az online térben állandóan ki van téve az árualapú logikának a reklámok formájában. Ez azt jelenti, hogy a legtöbb online idő reklámidő” (saját fordítás)

⁴¹ „A megfigyelési kapitalizmust működtetők mindent tudnak rólunk, miközben a műveleteiket úgy tervezték, hogy számunkra megismerhetetlenek legyenek” (saját fordítás)

lehetőségeinek rovására. Ez a piaci dominancia lehetővé teszi e vállalatok számára a politikai döntéshozatal befolyásolását is, mivel rendelkeznek a közpolitika számukra kedvező alakításához szükséges erőforrásokkal és befolyással. Az átláthatóság és az elszámoltathatóság hiánya e vállalatok működésében megnehezíti a demokratikus intézmények számára a hatékony felügyelet és szabályozás biztosítását. Pasquale szerint „*The more the black boxes of corporate practices in these areas are revealed, the more pressure will mount to change them (...) it would mean focusing less on trying to control the collection of data up front, and more on its use—how companies and governments are actually deploying it to make decisions*”⁴² (Pasquale, 2015, pp. 140-141).

A nagy technológiai vállalatok koncentrált gazdasági és politikai hatalma egyenlőtlen hatalmi dinamikát teremt ezen óriások és a szolgáltatásaikra támaszkodó felhasználók között. Az algoritmusok tervezőinek és alkalmazóinak előítéletei, elfogultságai és értékei kódolva lesznek a digitális életünket alakító rendszerekben, ami a faji, a nemi és a társadalmi-gazdasági státuszon alapuló diszkriminatív rendszerek kialakulásához vezethetnek.

Az egyének gondolatai, érzései és cselekedetei feletti befolyásnak ez a példátlan szintje mélyreható kérdéseket vet fel a szabad akarat természetéről és az önrendelkezés képességéről a digitális korban. E kihívások kezelése és az egyéni autonómia védelme érdekében elengedhetetlen a nagyobb átláthatóság, elszámoltathatóság és felhasználói ellenőrzés előmozdítása a digitális ökoszisztémán belül – különösen, hogy a mesterséges intelligencia és a gépi tanulási rendszerek elterjedésével ezek a problémák súlyosbodnak. Ez magában foglalhatja az adatvédelmi szabályok szigorítását, világos és tömör szolgáltatási feltételek előírását, valamint a felhasználók számára a személyes adataik és online tapasztalataik kezeléséhez szükséges eszközök biztosítását. A felügyeleti kapitalizmus e társadalmi és demokratikus következményeinek kezeléséhez elengedhetetlen a nagyobb átláthatóság és az elszámoltathatóság előmozdítása a digitális ökoszisztémában. Ez magában foglalhatja a trösztellenes szabályozás megerősítését, az AI-t használó rendszerek független ellenőrzésének előírását, valamint a felhasználók számára a személyes adataik feletti nagyobb kontrollt biztosító jogosítványokat – amikre elkezdődtek a globális kísérletek.

Ezen etikai normák és szabályozási keretek egyik elsődleges célja a technológiai vállalatok elszámoltathatóságának növelése kell, hogy legyen. Működésük átláthatatlansága és

⁴² „Minél inkább feltárulnak a vállalati gyakorlatok fekete dobozai ezeken a területeken, annál nagyobb lesz a nyomás, hogy változtassanak rajtuk. (...) kevésbé kellene összpontosítani az adatgyűjtés előzetes ellenőrzésére, és többet kellene foglalkozni az adatok felhasználásával - a vállalatok és a kormányok hogyan használják ténylegesen azokat a döntések meghozatalához” (saját fordítás)

az egyértelmű jogi és etikai határok hiánya lehetővé tette, hogy ezek a vállalatok kevés következménnyel járó kizsákmányoló gyakorlatot folytassanak. Ennek ellensúlyozására elengedhetetlen az átláthatóságot, felügyeletet és felelősséget előíró jogi és etikai követelmények megállapítása. Az etikai normák és szabályozási keretek másik kulcsfontosságú szempontja a felhasználói jogok, különösen a tájékoztatáshoz és a hozzájáruláshoz való jog előmozdítása kell, hogy legyen. A felhasználókat egyértelműen tájékoztatni kell arról, hogy az adataikat hogyan gyűjtik, dolgozzák fel és használják fel, és lehetőséget kell biztosítani számukra, hogy megadhatják vagy megtagadhatják a hozzájárulást ezekhez a gyakorlatokhoz. Ez magában foglalhatná egyértelmű és tömör szolgáltatási feltételek előírását, a vállalatok számára az adatgyűjtéshez és -felhasználáshoz való kifejezett hozzájárulás beszerzésének megkövetelését, valamint a felhasználók számára könnyen használható eszközök biztosítását adatvédelmi beállítások és adatmegosztási preferenciáik kezeléséhez. Nissenbaum rámutat, hogy „*a right to privacy as a right to control information about oneself*”⁴³ (Nissenbaum, 2011, p. 34). Azzal, hogy a felhasználóknak nagyobb ellenőrzést biztosítunk személyes adataik felett, és biztosítjuk tájékozott beleegyezésüket, elkezdhetjük újra egyensúlyba hozni a technológiai vállalatok és a szolgáltatásaikra támaszkodó egyének közötti hatalmi dinamikát. Az erős adatvédelmi jogszabályok kidolgozása és alkalmazása alapvető fontosságú a felhasználói jogok védelméhez. Az Európai Unió általános adatvédelmi rendelete (GDPR) olyan jogi keretet példáz, amelyek célja, hogy a felhasználóknak nagyobb ellenőrzést biztosítsanak személyes adataik felett, és a vállalatokat felelősségre vonhatóvá váljanak az adatkezelési gyakorlataikért.

Az elszámoltathatóságot, a felhasználói jogokat és az adatvédelmet előtérbe helyező etikai normák és szabályozási keretek létrehozásával és érvényesítésével olyan digitális ökoszisztémát hozhatunk létre, amely az egyének és a társadalom érdekeit szolgálja, nem pedig néhány domináns vállalat profitmotívumait.

Fel kell ismernünk, hogy a digitális technológiák, amelyek életünk nagy részét közvetítik, nem semleges eszközök, hanem inkább sajátos gazdasági szükségletek és hatalmi struktúrák termékei. Meg kell kérdőjeleznünk az e technológiák alapjául szolgáló gazdasági logikákat és értékeket, és olyan alternatív tervezési és alkalmazási módokat kell a helyükbe emelnünk, amelyek az emberi jólétet, a társadalmi igazságosságot és a demokratikus részvételt helyezik előtérbe – az előző fejezetben tárgyalt szabályozási kísérletek és ajánlások ezt célozzák.

⁴³ „*a magánélethez való jog, a személyre vonatkozó információk ellenőrzéséhez való jog*” (saját fordítás)

5.2.3 A mesterséges intelligencia és az automatizáció hatása a munka világra

A negyedik ipari forradalom (4IR) kezdetével a fejlett digitális technológiák, például a mesterséges intelligencia, a dolgok internete, a biotechnológia és a kvantumszámítástechnika elterjedése zökkenőmentesen integrálódott az ipari szférába, mélyreható változásokat idézve elő a vállalkozások működésében (Schwab, 2016). A fizikai, a digitális és a biológiai világ közötti határok elmosódása felgyorsította az operatív folyamatokat, elmosta a földrajzi korlátokat, és újradefiniálta a munkaerőpiac szerkezetét és elvárásait. Ezek a fejlesztések, miközben elhozták a példátlan hatékonyság és termelékenység korszakát, egyúttal bonyolult etikai kérdéseket is felvetnek a munka világában. A gyors technológiai fejlődés számos új tevékenységgel hozott létre, amelyek további bővítését elsősorban a munkahelyi hatékonyság és a termelékenység fokozására irányuló törekvés hajtja. Drucker (1999) szerint a hatékonyság és a termelékenység fokozása a tudásalapú gazdaságban az üzleti sikeresség sarokkövei, és az innováció, ha zökkenőmentesen beágyazódik a szervezeti működésbe, e sarokpontok eléréséhez vezető csatornaként szolgálhat.

5.2.3.1 A technológiai átalakulás és annak következményei a munkaerőpiacon

Az automatizálás, David H. Autor (2015) meghatározása szerint, azt a folyamatot foglalja magában, amikor a korábban emberi beavatkozást igénylő tevékenységeket a technológiai fejlődés révén, közvetlen emberi közreműködés nélkül végzik erre kifejlesztett eszközök. Ez lehetővé teszi az ipari szereplők számára, hogy az ismétlődő és időigényes feladatokból kivonják az emberi munkaerőt helyettesítve vagy kiegészítve azt, új lehetőségeket kínálva a munkavállalók számára, hogy a kreativitást, a kritikus gondolkodást és az érzelmi intelligenciát igénylő feladatokra összpontosíthassanak. A digitalizáció – megkülönböztetve a digitalizálástól – a digitális technológiák és adatok felhasználását jelenti az üzleti-termelési tevékenységek átalakítására, új vagy a meglévő üzleti folyamatok és ügyfélélmények létrehozására, a változó piaci igényeknek való megfelelés érdekében (Brennen & Kreiss, 2016). Ez az átalakulás elősegíti a döntéshozatal adatvezérelt megközelítését, lehetővé téve a hasznosítható információk pontos azonosítását, és elősegítve az agilisabb és alkalmazkodóbb üzleti modellek kialakítását.

Ahogy az egyes cégek integrálják a mesterséges intelligencia által vezérelt eszközöket, a valós idejű adatelemzés, a prediktív modellezés és a személyre szabott ismeretátadás előnyeit szerzik meg és alkalmazzák, átformálva a modern munkahelyeket. A mesterséges intelligencia munkahelyi integrációja új együttműködési modelleket hoz létre az emberi és a gépi szereplők között, lehetővé téve a cégek számára a termelékenység és a hatékonyság növelését (Daugherty

& Wilson, 2018). Ez a gyors technológiai változás számos kihívást hoz magával, például a munkavállalók új technológiákhoz való alkalmazkodásának kritikus kérdését is, illetve a konkrét munkafolyamatok automatizálásának eredményeképpen a megszűnő munkahelyek problematikáját (McAfee & Brynjolfsson, 2017).

A mesterséges intelligencia egyre kiterjedtebb alkalmazása egyszerre jelenti az emberi leleményesség forradalmát és a munkahelyről alkotott hagyományos elképzelések lényegi újragondolását, mivel gépek és az emberek közötti kölcsönhatás drasztikusan újralibrálódik, ami jelentős következményekkel jár a feladatokra, a funkciókra és az emberi munka lényegére nézve (Bessen, 2015). Az automatizálásra alkalmas feladatok és funkciók elemzése paradigmaváltást hoz a szervezeti stratégiákban, és elsősorban az ismétlődő, rutinszerű és a kialakult szabályokra vagy algoritmusokra támaszkodó feladatok azok, amelyek különösen veszélyeztetett helyzetbe kerülnek (Arntz et al., 2016). Az ilyen feladatok, amelyek az alapvető adatbeviteltől az olyan összetettebb szerepekig terjednek, mint a diagnosztikai elemzés vagy a pénzügyi tranzakciók, elsősorban a gépek által garantált hatékonyság, pontosság és költséghatékonyság miatt kerülnek átalakításra: a McKinsey Global Institute jelentése szerint a becslések szerint 2030-ig 400-800 millió ember kerülhet ki munkaköréből az automatizálás miatt, ami új foglalkozási szerepekre való átállást tesz szükségessé (Manyika et al., 2017, p. 11).

A technológiai fejlődés, miközben a szervezeteket páratlan hatékonysággal ruházza fel, egyúttal etikai kérdések szövevényes hálózatát is létrehozza, az emberi erőforrások fejlesztésének (HRD) területén is. Mivel a HRD elválaszthatatlanul kapcsolódik a szervezet legértékesebb tőkéjének – az emberi tőkének – a gondozásához és fejlesztéséhez, bármilyen innováció bevezetése ezen a területen szükségessé teszi az ezzel járó etikai következmények tudatosítását (Garavan et al., 2016). A HRD-ben az AI és a digitalizációval kapcsolatos etikai kihívások olyan kérdéseket foglalnak magukban, mint a mesterséges intelligencia által vezérelt toborzási (és kiválasztási) eszközök lehetséges elfogultságai (például a nemi vagy faji diszkrimináció), az alkalmazottak magánéletének eróziója a mindent átható digitális felügyelet miatt, valamint a digitalizált munkahely pszichológiai következményei, amelyek elmosásák a szakmai és a személyes szféra közötti határokat (Davenport et al., 2020).

A hatékonyság és az etikai megfontolások közötti egyensúly megteremtése bonyolult feladat. A növekvő termelékenység, amely gyakran az automatizálás és a mesterséges intelligencia alkalmazása révén alakul ki, arra csábíthatja a szervezeteket, hogy háttérbe szorítsák az emberi tényezőt. A hatékonyság azonban, ha egyedüli kritériumként tekintünk rá,

akaratlanul a munkavállalók marginalizálásához vezethet, ami potenciálisan munkahelyi egyenlőtlenségeket, a munkahelyi biztonság erózióját és a munkavállaló-munkáltató bizalmi kapcsolat romlását eredményezheti (Boudreau & Cascio, 2017). A Carroll (1991) által megfogalmazott vállalati etikai felelősségvállalás túlmutat a törvényeknek és szabályozásoknak való megfelelésen. Ez egy olyan önkéntes elkötelezettséget foglal magában, amely túllép a gazdasági megfontolásokon, és biztosítja, hogy a döntések, különösen azok, amelyek az emberi tőkét érintik, valós etikai elkötelezettségben gyökerezzenek – „*social responsibility can only become reality if more managers become moral instead of amoral or immoral*” (Carroll, 1991, p. 39). A szervezeteknek a humán erőforrás-fejlesztésen belüli innovációra való törekvésük során átfogó etikai értékeléseket, az érdekelt felekkel folytatott konzultációkat kell beépíteniük, és ami talán a legfontosabb, be kell fektetniük a munkavállalók folyamatos tovább- és átképzésébe, biztosítva, hogy a munkaerő egyetlen szegmense se szoruljon ki a munkaerőpiacról vagy veszítsen jogaiból a technológiai integráció szintjének növekedése nyomán (Aguinis & Glavas, 2012).

Egyre több cég ruház be úgynevezett „megmagyarázható mesterséges intelligencia” platformokba (Explainable artificial intelligence – XAI), ami olyan folyamatok és módszerek összességét jelöli, amelyek átláthatóvá és érthetővé téve az AI döntéshozatali folyamatait, lehetővé teszik az emberi felhasználók számára, hogy megértsék a gépi tanulási algoritmusok által létrehozott eredményeket és kimeneteket ezáltal növelve a bizalmat (Rudin, 2019).

Egy bizonyos szervezetméret felett megfontolandó külön etikai bizottságok létrehozása is, amelyek a humán erőforrás-fejlesztésben az AI és a digitalizáció következményeire összpontosítanak, biztosítva, hogy a technológiai integrációk összhangban legyenek az etikai irányelvekkel és az emberi jóléttel. Az ilyen intézkedések magukban foglalhatják a szakképzettség növelésére és átképzésére irányuló kezdeményezéseket, amelyek biztosítják, hogy az automatizálás vagy a digitalizáció által kiszorított munkavállalók rendelkezzenek az új szerepkörökbe való átmenethez szükséges készségekkel (World Economic Forum, 2020). Ezen túlmenően a mentális egészség megőrzését és javítását, valamint az általános munkavállalói jóllétet támogató programokkal kezelhetők a gyors munkahelyi átalakulások pszichológiai következményei. A gyakorlatban az etikai megfontolások beépítése a termékekbe magában foglalja az irányelvek kialakítását bizonyos terméktípusok etikus tervezéséhez és fejlesztéséhez, erkölcsi nézőpontok bevonását a termékfejlesztésbe, az alkalmazottak képzését és tájékoztatását az etikai kérdésekről.

5.2.4 A mesterséges intelligencia alkalmazásának etikai vonatkozásai az oktatásban

Egy korábbi tanulmányomban (Szilágyi, 2024b, pp.211-213) már írtam arról, hogy az AI-vezérelt oktatási innovációk jelentős változást jelentenek az oktatásban, mivel személyre szabott, interaktív és hatékony tanulási lehetőségeket kínálnak. A mesterséges intelligencia egyre inkább beépül az oktatási folyamatokba, radikális változások és átalakulások lehetőségét kínálva, egyúttal olyan összetett etikai problémák során vetve fel, mint a magánélet védelme, a beleegyezés korlátai és a digitális szakadék kérdése.

A mesterséges intelligencia által vezérelt tanulási innovációk esetében – ahogy az előzőekben már említett területeken is – a magánélet védelme kiemelt problémaként jelentkezik, mivel a hatékony AI rendszerek nagy mennyiségű adatot igényelnek, beleértve a tanulók tanulási szokásairól, személyes jellemzőiről és teljesítményéről szóló érzékeny információkat, ami számos aggályt vet fel azzal kapcsolatban, hogy ezeket az információkat hogyan tárolják, ki fér hozzá, és hogyan használják fel. Ezek az új tanulási eszközök algoritmusok segítségével elemzik az egyes tanulók teljesítményét, preferenciáit és viselkedését, hogy személyre szabott visszajelzést és tanulási utakat és módszereket kínáljanak. A rendszerek részletes információkat képesek nyomon követni, például azt, hogy egy tanuló mennyi időt tölt egy adott feladattal, milyen a sikerességi aránya és milyenek az interakciós mintái. Bár a tanulási élmény fokozása szempontjából előnyös, ez az adatintenzív megközelítés jelentős adatvédelmi kockázatokat rejt magában. Ezt az aggodalmat jól szemléltetik az online tesztelésben egyre szélesebb körben alkalmazott mesterséges intelligencia alapú ellenőrző eszközök, amelyek a csalás megelőzése érdekében figyelemmel kísérik a diákok viselkedését a tesztek során. Ezek az eszközök gyakran hozzáférést igényelnek a tanuló webkamerájához és mikrofonjához, elemzik a fizikai és környezeti adatokat, és akár a tanuló szemmozgását és billentyűleütéseit is megfigyelhetik. Itt elmosódik a határ az etikus korlátok között végzett szükséges felügyelet és egy, a privát szférát sértő invazív megfigyelés között, ami komoly kérdéseket vet fel azzal kapcsolatban, hogy az oktatás integritása érdekében milyen mértékben lehet veszélyeztetni a diákok magánéletét.

Ahogy a mesterséges intelligencia tanulási környezetekben való alkalmazása egyre elterjedtebbé válik, kiemeli annak szükségességét, hogy a tájékozott beleegyezést alapvető etikai követelménynek és jogi kötelezettségnek tekintsük a legtöbb területen, így az oktatásban is. Ez a koncepció a digitális terekben, különösen a mesterséges intelligencia és a nagy mennyiségű adat esetében bonyolulttá válik, mivel e technológiák összetettsége és átláthatatlansága megnehezítheti a diákok és gondviselőik számára, hogy teljes mértékben

megértsék, mihez járulnak hozzá. Az információk típusa, amelyekhez a diákok vagy gyámjaik hozzájárulását kérik, a személyes azonosítóktól (mint a név és személyi igazolványszám) az olyan érzékeny információkig terjedhet, mint a tanulási szokások, fogyatékoságok vagy más egyéni jellemzők, amelyek segítenek a tanulási élmény személyre szabásában. Ideális esetben ez a hozzájárulás kifejezett, tájékozott, és ésszerű időközönként vagy az adatok felhasználásának jelentős változása esetén megújítandó. A kihívás az, ahogy Joseph W. Jerome (Jerome, 2017) megjegyzi, hogy a *Big Data* korában az adatgyűjtési gyakorlatok összetettsége és átláthatatlansága megnehezítheti az egyének számára, hogy megértsék és értelmesen hozzájáruljanak ezekhez a gyakorlatokhoz. Előfordulhat, hogy a tanulók és gondviselőik nem értik teljesen a hozzájárulásuk hosszú távú következményeit, illetve azt, hogy adataik a közvetlen oktatási kontextuson túl hogyan használhatók fel (Jerome, 2014, pp. 234-239). Ez kapcsolódik Franklin G. Miller és Alan Wertheimer érveihöz, akik hangsúlyozzák, hogy a tájékozott beleegyezésnek többnek kell lennie, mint pusztán formalitásnak; folyamatos kommunikációs folyamatnak kell lennie, amelyben az egyének világosan megértik, hogy mihez adják a beleegyezésüket. A beleegyezés nem csak az érvényes hozzájárulásról szól, hanem arról is, hogyan védi és mozdítja elő az egyén jólétét, autonómiáját és támogatja az önrendelkezési jogát, vagyis a beleegyezés, mint aktus erkölcsi következményekkel jár, transzformatív jelleggel bír (Miller & Wertheimer, 2010). Ezen elképzelések alkalmazása a mesterséges intelligenciára az oktatásban azt jelenti, hogy az oktatási intézmények és a technológiai vállalatok folyamatosan átláthatóvá teszik adatkezelési gyakorlatukat, és folyamatos párbeszédet folytatnak a diákokkal és a szüleikkel. Ez lehetővé teszi, a mélyebb és tájékozott megértést, az AI potenciális előnyeiről és kockázatairól, ami az egyéni autonómia védelmén túl, elősegíti az önrendelkezési jog gyakorlását is azáltal, hogy képesek lesznek informált döntéseket hozni arról, milyen mértékben kívánnak részt venni az AI-alapú oktatási programokban, használni új technológiákat.

Ezek az aggodalmak visszhangzanak Beate Roessler *The Value of Privacy* (Roessler, 2005) című írásában, ahol a magánélet árnyaltabb értelmezése mellett érvel, és azt olyan alapvető emberi jogként mutatja be, amely magában foglalja az egyének személyes adataik feletti ellenőrzést. A mesterséges intelligenciára az oktatásban alkalmazva ez a szemlélet az adatgyűjtési, tárolási és felhasználási mechanizmusok gondos átvizsgálását követeli meg, különösen figyelembe véve a legfiatalabb generációk kiszolgáltatottságát az oktatási környezetben. E kérdések fényében világossá válik, hogy miközben a mesterséges intelligencia által vezérelt tanulási innovációk számos előnyt ígérnek, gondosan figyelembe kell venniük az

oktatásban résztvevők magánéletét. Ez az oktatásban az adatvédelemmel kapcsolatos szigorú szabályozást és a tanulók jogait előtérbe helyező mesterséges intelligenciát használó rendszerek etikus tervezését teszi szükségessé. A tanulóknak és azok szüleinek tájékozottnak kell lenniük és meg kell érteniük, hogy a mesterséges intelligenciát használó eszközök és programok milyen adatokat gyűjtenek az emberekről és azokat hogyan használják fel, és ha úgy döntenek, miként rendelkezhetnek azok felhasználásának módjairól és korlátaival.

5.2.4.1 A mesterséges intelligencia elfogultsága és a digitális szakadék problematikája

Timnit Gebru és Joy Buolamwini egy kutatás során vizsgálta az AI-algoritmusokban rejlő elfogultságokat, és megállapították, hogy az AI-algoritmusok, különösen a gépi tanulásban használtak, nem semlegesek, hanem inkább tükrözik a rájuk képzett adatok és a készítőik nézőpontjainak elfogultságait, ami olyan AI-rendszereket eredményezhet, amelyek állandósítják és felerősítik a faji és nemi előítéleteket. Az egyik kritikus példa ennek illusztrálására, az arcfelismerő technológia, aminek kapcsán rámutatnak, hogy ezeket a rendszereket gyakran olyan adathalmazokon képzik ki, amelyek túlnyomórészt fehér férfi arcokból állnak. Ennek eredményeképpen általában rosszul teljesítenek a színesbőrűek és nők arcának felismerésekor, ami jelentős eltéréseket eredményez a pontosságukban. Ezen problémák megoldása érdekében nagyobb átláthatóságot és elszámoltathatóságot szorgalmaznak, olyan módszerek kidolgozását, amelyek lehetővé teszik az AI-algoritmusok ellenőrzését és vizsgálatát, a döntéshozatali folyamatok átláthatóságát és érthetőségének javítását, valamint az előítéleteknek minősíthető gyakorlatok azonosítását és korrigálását: „*We define accountability as reporting algorithmic performance on demographic and phenotypic subgroups and actively working to close performance gaps where they arise. Algorithmic transparency and accountability extend beyond technical reports and should include mechanisms for consent and redress*”⁴⁴ (Buolamwini & Gebru, 2018, p. 12)

Ruha Benjamin a *Race after technology: Abolitionist tools for the new Jim code* (Benjamin, 2019) című könyvében, amelyben kritikusan vizsgálja a faj, a technológia és a társadalom metszéspontját – kiemeli, hogy a technológia, beleértve a mesterséges intelligenciát is, nem semleges, hanem olyan előítéletekbe van beágyazva, amelyek állandósíthatják és súlyosbíthatják a meglévő társadalmi egyenlőtlenségeket, és a faji alapon elfogult

⁴⁴ „Az elszámoltathatóságot úgy határozzuk meg, mint az algoritmusok teljesítményének folyamatos felülvizsgálatát demográfiai és fenotípusos alcsoportokra vonatkozóan, valamint aktív munkát a felmerülő teljesítménybeli különbségek megszüntetésére. Az algoritmikus átláthatóság és elszámoltathatóság túlmutat a technikai jelentéseken, és magában kellene foglalnia a beleegyezés és a jogorvoslat mechanizmusait is.” (saját fordítás)

technológiákat a Jim Crow-törvényekhez hasonlítja, amelyek megalapozták és megszilárdították a szegregációt az Egyesült Államokban. A semlegesség és objektivitás ürügyén a technológia megerősítheti a faji hierarchiákat – áll Benjamin írásában. A könyv egyik fő érve, hogy a technológia egyfajta „kódolt egyenlőtlenség” eszköze lehet, ahol a faji előítéletek kódolva vannak a technológiai rendszerekbe. Ez többféleképpen történhet, például a mesterséges intelligenciát használó rendszerek képzéséhez használt elfogultnak minősíthető adathalmazok, vagy a technológiák olyan módon történő megtervezése és megvalósítása révén, amely hátrányos helyzetbe hoz bizonyos faji vagy etnikai csoportokat. Benjamin azonban azt is hangsúlyozza, hogy a technológia mindezek mellett a társadalmi igazságosság eszközeként is használható és a technológia egyfajta „abolicionista” megközelítése mellett száll síkra, amely magában foglalja az ilyen előítéletek aktív lebontását, és a technológia felhasználását a társadalmi egyenlőtlenségek kiküszöbölésére.

A mesterséges intelligencia elterjedése az oktatásban magában hordozza a kockázatot, hogy elmélyíti a meglévő társadalmi különbségeket. A digitális szakadék, amely a kommunikációs és információs technológiákhoz való hozzáférés terén mutatkozó egyenlőtlenségeket írja le, azon veszélyeit vetíti előre az AI-alapú tanulási módszerek széles körű alkalmazásának, amely a társadalmi csoportok közötti távolságokra is kihatnak. Annak ellenére, hogy az AI számos lehetőséget kínál a pedagógiai gyakorlatok és tanulási folyamatok fejlesztésére, megfelelő körültekintés hiányában paradox módon felerősítheti azokat a hátrányokat, amelyeket eredetileg enyhíteni hivatott (Bentley et al., 2024). Ezért kulcsfontosságú, hogy az oktatási szektorban az AI bevezetését átgondolt, a társadalmi egyenlőséget szem előtt tartó stratégiák mentén valósítsák meg, mivel a technológiához való hozzáférés döntő szerepet játszik az eltérő szociokulturális háttérű csoportok tanulási lehetőségeinek és eredményeinek alakításában. Az, hogy a tanulók képesek-e hozzáférni és használni a mesterséges intelligencia alapú oktatási eszközöket, számos tényezőtől függ, többek között a megbízható internetkapcsolattól, a megfelelő eszközök birtoklásától és a digitális írástudás szintjétől. Ezért a mesterséges intelligencia előnyei a tanulásban nem egyenletesen oszlanak meg, gyakran előnyben részesítve az erőforrásokhoz jobban hozzáférőket, ezáltal súlyosbítva a társadalmi-gazdasági egyenlőtlenségeket. A neves politológus Pippa Norris *Digital Divide: Civic Engagement, Information Poverty and the Internet World-Wide* című könyvében (Norris, 2001) a digitális szakadékot többdimenziós kérdésként jellemzi, amely túlmutat az IKT-hoz való puszta hozzáféréseken: azt állítja, hogy még azok között is, akik hozzáférnek a technológiához, létezik egy megosztottság azok között, akik hatékonyan tudják

használni azt a társadalmi részvételre, és azok között, akik nem tudják használni, amit ő „demokratikus szakadéknak” nevez. A digitalizáció egy korábbi szakaszában fogalmazva meg kételyeit, az internet terjedése kapcsán merül fel benne a tartós eltérések kialakulásának és konzerválódásának lehetősége:

„Projections of future developments remain uncertain and hazardous, surrounded by health warnings, yet at the same time it is even more important that we seek to understand the underlying reasons for patterns of Internet access and use at this stage in the diffusion process, even imperfectly, before the initial inequalities rigidify into a virtual Berlin Wall dividing the information-rich and poor, within and between societies”⁴⁵(Norris, 2001, p. 240).

Ez a probléma a mesterséges intelligenciát használó rendszerek és eszközök robbanásszerű terjedésével csak tovább mélyül, nem pusztán a kétezres évek elején Alexander Bard és Jan Söderqvist által felvetett *netokrácia* lehetséges elkülönülésének szintjén (Bard & Söderqvist, 2022), hanem a társadalmi struktúra radikális megváltozásához is elvezethet (ahogy korábban a munkaerőpiaci hatásokról már tárgyaltam).

A digitális szakadék kérdése tehát jelentős etikai következményekkel jár a mesterséges intelligencia tanulásban való alkalmazása szempontjából. A mesterséges intelligencia oktatásba való beépítésére irányuló erőfeszítéseknek tisztában kell lenniük ezekkel az egyenlőtlenségekkel, és törekedniük kell arra, hogy a mesterséges intelligencia az oktatási egyenlőség növelésének eszköze legyen, ne pedig a meglévő egyenlőtlenségek súlyosbításának eszköze. E kérdések kezelésére többirányú megközelítés lehet indokolt, amely magában foglalja a politikai szabályozást, a technikai fejlesztéseket és az oktatási intézkedéseket. A jövőbeni kutatásoknak az etikusan tervezett mesterséges intelligenciát használó rendszerek fejlesztésére, a hatékony beleegyezési mechanizmusok biztosításának módjainak feltárására, valamint a digitális szakadék áthidalására irányuló stratégiák kidolgozására kell összpontosítani az AI-oktatásban. Ezen túlmenően az újonnan megjelenő mesterséges intelligencia-technológiák etikai következményeinek az oktatásban való kutatásának folyamatos erőfeszítéseket kell jelentenie a fejlesztésükkel és alkalmazásukkal párhuzamosan. Ahogy ezek a technológiák fejlődnek, valószínűleg új etikai kihívások merülnek majd fel, ami szükségessé teszi a mesterséges intelligencia tanulásban való etikus felhasználását biztosító megközelítéseink folyamatos vizsgálatát és kiigazítását (Szilágyi, 2024b, p. 213).

⁴⁵ „A jövőbeli fejlemények előrejelzése bizonytalan és kockázatos marad, figyelmeztető jelek kíséretében, de még fontosabb, hogy megértsük az internetes hozzáférés és használat mintáinak mozgatórugóit ebben a diffúziós folyamatban, még ha tökéletlenül is, mielőtt az eredeti egyenlőtlenségek szigorú szétválása egy virtuális berlini falat eredményezne, amely elválasztja az információban gazdag és szegény rétegeket, mind a társadalmon belül, mind a társadalmak között.” (saját fordítás)

VI. A felelős döntéshozatali mátrix, mint az etikai értékelés keretrendszere

6.1 A félelem konstruktív becsatornázása

Az egyik fő kockázata a mesterséges intelligencia alapú rendszereknek, hogy az ember elveszíti a kontrollt, cselekvő- és beavatkozási képessége korlátozódik. A felvetésem arra irányul, hogy ha gondosan mérlegeljük a technológia által jelentett kockázatokat és kihívásokat, és félelmeinket etikai döntéshozatali útmutatóként használjuk, akkor technológiai fejlődés előnyeit úgy használhatjuk ki, hogy közben minimalizáljuk a lehetséges károkat.

Hans Jonas korábban már tárgyalt felelősségetikája, különösen a *félelem heurisztikája* koncepció, egy működőképes keretet kínál ezeknek a kihívásoknak a kezelésére. Jonas szerint a jövővel kapcsolatos félelmeink fontos szerepet játszanak a technológiai fejlesztésekkel kapcsolatos etikai döntéshozatalban. A félelem egy hasznos jelzés, amely segít a kockázatok felmérésében és a felelős cselekvésben, ugyanakkor a félelem nem vezethet a technológiai fejlődés elutasításához, hanem a tudatos, értékvezérelt innovációt kell ösztönöznie. Kiemelt jelentőségűvé válik a félelem konstruktív csatornázása: a félelem heurisztikáját úgy kell alkalmazni, hogy az konstruktív cselekvésre ösztönözzön, ne bénító hatású legyen (Jonas, 1984). Az AI fejlesztésére és alkalmazására vonatkozóan kidolgozott irányelveknek tükrözniük kell a félelem heurisztikájából fakadó elővigyázatossági elvet, de ez nem vezethet technológiaellenességhez, hanem a felelős innováció motorjává kell, hogy váljon. Ez a megközelítés kényes egyensúlyt igényel az új technológiákban rejlő lehetőségek felkarolása, valamint az óvatosság és a korlátok és veszélyek szem előtt tartása között. Végző soron a cél az, hogy félelmeinket konstruktív módon irányítsuk, felhasználva azokat a felelős innováció ösztönzésére és egy olyan jövő megteremtésére, amelyben a technológia az emberiség igényeit és értékeit szolgálja.

Hans Jonas szerint a félelem nem egy negatív érzelm, amelyet el kell kerülni, hanem inkább egy olyan fontos jel, amely segíthet a kockázatok felmérésében és a bizonytalansággal szemben a felelősségteljes cselekvésben. A technológiai fejlődéssel összefüggésben a félelem „heurisztikaként”, problémamegoldó stratégiaként szolgálhat, amely vezetheti a döntéshozatalunkat, és segít előre látni és mérsékelni a potenciális veszélyeket: „*We know the thing at stake only when we know that it is at stake*”⁴⁶ – írja (Jonas, 1984, p. 27). Ha odafigyelünk félelmeinkre, és azokat etikai reflexióink kiindulópontjaként használjuk, óvatosabb és felelősségteljesebb megközelítést alakíthatunk ki az innovációval kapcsolatban.

⁴⁶ „Csak akkor tudjuk, mi forog kockán, ha tudjuk, hogy kockán forog” (saját fordítás)

Jonas elméletének középpontjában az a gondolat áll, hogy a félelem kulcsszerepet játszhat és kell játszania a technológiai fejlesztésekkel kapcsolatos etikai döntéshozatalban. A hagyományos etikai keretek, amelyek gyakran az egyéni cselekvésekre és azok közvetlen következményeire összpontosítanak, nem alkalmasak a technológiai fejlődés következményeinek kezelésére.

Bár a félelem értékes stratégiai eszköz lehet a kockázatok felmérésében, fontos felismerni, hogy a félelem önmagában nem elegendő a felelős innováció irányításához. Ehelyett a félelem heurisztikáját konstruktív módon a tudatos, értékközpontú innováció felé kell terelni, amely az etikai megfontolásokat helyezi előtérbe (Floridi et al., 2018). A mesterséges intelligencia fejlesztésével összefüggésben ez annak biztosítását jelenti, hogy az etikai elvek és értékek a tervezés, fejlesztés és bevezetés teljes folyamatába beágyazódjanak. Ehhez proaktív és folyamatos elkötelezettségre van szükség az etikai reflexió és elemzés iránt, valamint olyan szilárd etikai keretek és iránymutatások kidolgozására, amelyek segíthetnek az AI-kutatás és -fejlesztés irányának irányításában (Jobin et al., 2019). A mesterséges intelligencia fejlesztésével összefüggésben a félelem heurisztikája azt sugallja, hogy az óvatosság oldalára kell állnunk, és a biztonsági és védelmi szempontokat kell előtérbe helyeznünk, még akkor is, ha ez a technológiai fejlődés ütemének lassítását vagy korlátozását jelenti. Jonas úgy fogalmaz, hogy:

„There are times when the drive needs moral encouragement, when hope and daring rather than fear and caution should lead. Ours is not one of them. In the headlong rush, the perils of excess become uppermost. This necessitates an ethical emphasis which, we hope, is as temporary as the condition it has to counteract. But there is also a timeless precedence of »thou shall not« over »thou shalt« in ethics”⁴⁷
(Jonas, 1984, p. 203).

A félelem által inspirált gondolkodás arra ösztönzi a szakembereket, hogy az etikai szempontokat már a tervezés és fejlesztés korai fázisaiban integrálják a folyamatokba. Ez jelentheti például azt, hogy az AI -rendszereket összhangba hozzák bizonyos etikai normákkal, gondoskodnak az átlátható és számonkérhető működésről, valamint támogatják a technológiáról folyó társadalmi diskurzust. A bizonytalan kockázatok felmerülése esetén a félelem heurisztikája azt diktálja, hogy a helyezzük előtérbe a biztonságot, aminek az eredményeképpen például a technológiai cégek bevezethetnek szigorúbb biztonsági

⁴⁷ „Vannak idők, amikor a hajtóerőnek erkölcsi bátorításra van szüksége, amikor a reménynek és a merészségnek kell vezetnie, nem pedig a félelemnek és az óvatosságnak. A miénk nem ilyen. A fejevesztett rohanásban a túlzások veszélyei kerülnek előtérbe. Ez olyan etikai hangsúlyt tesz szükségessé, amely reményeink szerint éppoly átmeneti, mint az az állapot, amelyet ellensúlyoznia kell. De az etikában a »ne tedd« időtlen elsőbbséget élvez a »tedd« felett is” (saját fordítás)

előírásokat, megkövetelhetnek alaposabb tesztelési eljárásokat, esetleg lelassíthatnak vagy akár felfüggeszthetnek bizonyos fejlesztéseket mindaddig, amíg a kockázatok pontosabban fel nem mérhetők. Az esetleges negatív következményektől való félelem arra is motiválhatja a fejlesztőket, hogy innovatívabb és biztonságosabb megoldásokat keressenek.

6.2 A félelem heurisztikája, mint a felelős döntéshozatali mátrix alapelve

Ennek a fejezetnek a központi állítása az, hogy a félelem heurisztikája, ahogyan Hans Jonas javasolta, értékes eszközként szolgálhat a felelős innováció ösztönzésére az olyan potenciálisan destruktív technológiákkal kapcsolatban, mint a mesterséges intelligencia és kialakítható egy olyan keretrendszer, amely segítségével az új fejlesztések felelősségtikai értékelését el lehet végezni. A következőkben egy ilyen általam tervezett etikai döntéshozatali keretrendszer bevezetésére teszek kísérletet, amelyet a továbbiakban *felelős döntéshozatali mátrixnak* nevezek. Ez a mátrix egy komplex, széleskörűen alkalmazható keretet kínál a mesterséges intelligencia által támasztott etikai kihívások megközelítéséhez és az egyes kihívásokra adandó válaszok megfogalmazásához. Négy olyan általam elkülönített alapelvet foglal magában, amelyeknek a mesterséges intelligencia fejlesztésével és alkalmazásával kapcsolatos döntéshozatali és cselekvési folyamatainkat irányíthatják a felelősségtikai elvek alapján: [1] a hosszú távú következmények figyelembevétele (*impvizibilis temporális hatás*), az [2] elővigyázatosság (*prekaucionalitás*) elve, a [3] globális hatások (*globális hatás elv*) figyelembevétele és a [4] jövő generációk iránti felelősség (*intergenerációs hatások elve*).

Az első alapelv, az *impvizibilis temporális hatások* figyelembevételének elve, azon a felismerésen alapul, hogy az AI rendszerek hosszú távon kiszámíthatatlan módon befolyásolhatják és alakíthatják az emberi életvitelt, a gazdasági és politikai struktúrákat, valamint a társadalmat. Ez alapján arra kényszerülünk a döntéshozatali folyamatainkban, hogy az azonnali előnyökön vagy kockázatokon túlmutatva mérlegeljük, hogy a mesterséges intelligencia milyen ontológiai kérdéseket vet fel, vagy gyakorlati módon: miként befolyásolhatja az emberi létezés feltételeit az elkövetkező évtizedekben és évszázadokban. Ez további kérdéseket implicál az ember és az AI közötti lehetséges interakciók fejlődésével, a munkaerőpiac átalakulásával és a társadalmi struktúrák megváltozásával kapcsolatban.

A *prekaucionalitás elve* a mátrix második eleme, a kockázatcsökkentés kritikus fontosságát hangsúlyozza a mesterséges intelligencia fejlesztésében, középpontba a mérlegelést állítva (vezérgondolata a már korábban említett félelem heurisztikája). Ez az elv komplex kockázatértékelésre, biztosítékok bevezetésére és vészhelyzeti tervek kidolgozására ösztönöz.

Felhívja a figyelmet arra, hogy az olyan több társadalmi alrendszerre kiterjedő transzformatív technológiákkal kapcsolatban, mint a mesterséges intelligencia, egy olyan inherens elővigyázatosság szükséges, amely a felelős innováció alapját képezi, biztosítva, hogy bármilyen fejlesztés átgondoltan és a lehetséges negatív következmények kellő figyelembevételével történhessen.

A harmadik *a globális hatás elv* a lokálitáson, valamint a nemzeti kereteken és határokon is túlmutató perspektíva felvázolását kívánja meg. Abból a felismerésből indul ki, hogy amennyiben a mesterséges intelligenciát alkalmazó rendszerek egyszer bevezetésre és alkalmazásra kerülnek, olyan következményekkel járhatnak, amelyek túlmutatnak a földrajzi és kulturális határokon. Ez az elv arra kényszerít minket, hogy megvizsgáljuk, miként hathat az AI az eltérő kulturális háttérrel rendelkező társadalmakra, felmérjük a mesterséges intelligencia környezeti hatásait, figyelembe véve, hogy potenciálisan enyhíti vagy súlyosbítja a mesterséges intelligencia és bármilyen alkalmazása az egyenlőtlenségeket.

A mátrix negyedik alapelvét az *intergenerációs hatások elve* képezi, amely a ma kifejlesztett mesterséges intelligenciát használó rendszerek jövő generációira gyakorolt lehetséges hatásaira összpontosít, magába foglalva több, a felelősségetikai diskurzusból ismert elvet. Az intergenerációs hatások elve fókuszában a következő nemzedékekkel szembeni felelősség gondolata áll: olyan technológiát érintő döntéseket szabad csak meghoznunk a jelenben, amelyek kapcsán felmérjük, hogyan befolyásolhatja a jövő nemzedékek lehetőségeit és életminőségét. Mérlegelnünk kell, miként biztosíthatjuk azt, hogy a mesterséges intelligencia fejlesztése megőrizze vagy növelje a jövő generációinak autonómiáját és cselekvőképességét, egyensúlyt teremtve a mesterséges intelligenciában rejlő izgalmas lehetőségek és a lehetséges kockázatok és nem szándékolt következmények értékelése között.

A felelős döntéshozatali mátrix kijelöli az alkalmazásának a korlátait is, rávilágít az adaptív tervezési és irányítási struktúrák és politikák szükségességére, amelyek képesek együtt változni a gyorsan fejlődő AI technológiákkal együtt. A mátrix alkalmazása kiemeli az olyan inkluzív döntéshozatali folyamatok fontosságát, amelyekben különböző érdekelt felek vesznek részt, a tudósoktól-fejlesztőktől és a politikai döntéshozóktól kezdve, az alkalmazott etikai szakértőkön át a potenciálisan érintett közösségek képviselőikig.

A felelős döntéshozatali mátrix alapján meghatározhatunk egy összesített mutatót, a *reszponzibilitási indexet*, amely az adott innováció etikai megfelelőségét és felelősségvállalási szintjét számszerűsíti. Ez az index az innováció különböző etikai dimenziókban mutatott teljesítményének átlagos értékeléseiből kerül kiszámításra. *A reszponzibilitási index az etikai*

dimenziókhoz tartozó alapelvek alapján kapott pontszámok átlagából adódó mutató, amely azt jelzi, hogy az adott innováció milyen mértékben felel meg a felelős tervezés etikai követelményeinek, beleértve a hosszú távú hatások, a kockázatkezelés, a globális hatások és a jövő generációk iránti felelősség szempontjait.

6.2.1 Értékelési rendszer

A felelős döntéshozatali mátrix alkalmazása egy interszubjektív értékelési folyamat során történik, amelynek az a célja, hogy több személy vagy csoport bevonásával egy átfogó és kiegyensúlyozott értékelést tudjunk kialakítani az adott innováció etikai dimenzióiról. A folyamat első lépése az értékelők kiválasztása, amely csoport kialakításakor törekedni kell a sokszínűsége, így az érintettek bevonása mellett a releváns szakterületek képviselőire is. Ezek alapján az interszubjektív csoportba be kell vonnunk az innovációs folyamatban közvetlenül érintett fejlesztőket, akik a fejlesztés céljaival és a technikai részletekkel kapcsolatosan mélyreható ismeretekkel rendelkeznek. Emellett ajánlott bevonni az innováció célcsoportját, és általánosságban a fejlesztés által érintett közösségek képviselőit (ezek meghatározásában segítségünkre lehet a stakeholder-analízis módszertana), mivel értékes betekintést adhatnak az innováció potenciális hatásaival kapcsolatban. Bevonhatjuk továbbá a szabályozás terén érintett hatóságok vagy civil szervezetek képviselőit, valamint nem opcionálisan olyan alkalmazott etikai szakértőket, akik jártasak az etikai elméletek és azok gyakorlati alkalmazása terén. Általános elvként elmondható, hogy az értékelők kiválasztásánál figyelembe kell venni a szakmai kompetenciákat, a függetlenséget és az etikus és felelős innováció iránti elkötelezettséget.

Az interszubjektív értékelési folyamat következő szakasza az egyéni értékelések elkészítése, amely szakaszban az összes értékelő önállóan, a saját szakterületének és tapasztalatainak megfelelően a meghatározott módszertant követve, pontszámokkal értékeli az innovációt a felelős döntéshozatali mátrix etikai dimenzióinak és kérdéseinek mentén. Az egyéni értékelések részét képezik az értékelők részletes indoklásai, amelyeket a pontszámokhoz fűznek és az összegzést követően, egy komparatív elemzés során próbálnak közös álláspontot kialakítani a felhasználásukkal. A következő táblázatban a döntéshozatali mátrix felépítését láthatjuk, az egyes dimenziókhoz kapcsolódó értékelőkérdésekkel.

8. táblázat. Felelős döntéshozatali mátrix kérdései (saját szerkesztés)

Etikai dimenziók	(I.) Imprevizibilis temporális hatások elve	(II.) Prekacionalitás elve	(III.) Globális hatás-elv	(IV.) Intergenerációs hatások elve
(A) Emberi élet és méltóság tisztelete	<p>1. Milyen mértékben képes ez az innováció hosszú távon megőrizni és támogatni az egyének méltóságát és alapvető jogait, figyelembe véve a változó társadalmi és technológiai környezetet?</p> <p>2. Mennyire valószínű, hogy ez az innováció előre nem látható következményekkel járhat, amelyek hosszú távon veszélyeztetik az emberi életet és méltóságot?</p>	<p>1. Milyen mértékben rendelkezik ez az innováció biztosítékokkal az emberi élet és méltóság védelmére a fejlesztés lezárultát követően?</p> <p>2. Mennyire biztosítja ez az innováció az egyének autonómiájának és cselekvőképességének hosszú távú megőrzését?</p>	<p>1. Milyen mértékben tartja ez az innováció tiszteletben az eltérő kultúrák emberi életre és méltóságra vonatkozó értékeit és nézeteit?</p> <p>2. Mennyire zárja ki ez az innováció annak lehetőségét, hogy negatív hatással legyen a marginalizált vagy kiszolgáltatott népcsoportokra világszerte?</p>	<p>1. Milyen mértékben biztosítja ez az innováció, hogy megőrizze vagy javítsa a jövő generációk lehetőségeit a méltósággal való életre?</p> <p>2. Mennyire zárja ki ez az innováció annak a lehetőségét, hogy olyan helyzeteket eredményezzen, amelyek korlátozzák a jövő generációk jogait vagy szabadságát?</p>
(B) Magánélet és adatvédelem	<p>1. Mennyire zárja ki ez az innováció a személyes adatokkal való visszaélés jövőbeli lehetőségét?</p> <p>2. Mennyire biztosítja ez az innováció a személyes adatok védelmét szolgáló intézkedések hosszú távú fenntarthatóságát és fejlesztését?</p>	<p>1. Milyen mértékben képes ez az innováció proaktívan kezelni az adatvédelmi kockázatokat, biztosítva a folyamatos védelmet a fejlődő technológiai környezetben?</p> <p>2. Mennyire biztosítja ez az innováció az egyének számára a személyes adataik feletti hosszú távú ellenőrzés megőrzését?</p>	<p>1. Mennyire tartja tiszteletben ez az innováció a különböző országok adatvédelmi normáit és elvárásait?</p> <p>2. Mennyire biztosítja ez az innováció, hogy ne fokozza a magánélethez való jog vagy az adatvédelem terén fennálló globális egyenlőtlenségeket?</p>	<p>1. Milyen mértékben védi ez az innováció a jövő generációinak a magánélethez való jogát?</p> <p>2. Mennyire zárja ki ez az innováció annak a lehetőségét, hogy olyan adathalmazt hozzon létre, amely veszélyezteti az egyének jövőbeli magánéletét?</p>
(C) Diszkrimináció és előítéletmentesség	<p>1. Milyen mértékben képes ez az innováció hosszú távon csökkenteni az előítéleteket és a diszkriminációt, elősegítve az esélyegyenlőséget?</p> <p>2. Mennyire segíti ez az innováció a társadalmi kohézió fenntartását, megelőzve a polarizáció növekedését?</p>	<p>1. Milyen mértékben képes ez az innováció a potenciálisan benne rejlő előítéletek feltárására és mérséklésére?</p> <p>2. Mennyire biztosítja ez az innováció a társadalmi befogadás elősegítését?</p>	<p>1. Milyen mértékben tartja tiszteletben ez az innováció a különböző kultúrák és közösségek sokszínűségét?</p> <p>2. Milyen mértékben veszi figyelembe és kezeli ez az innováció a lehetséges kulturális különbségekből adódó eltérő hatásokat, biztosítva, hogy ne gyakoroljon aránytalan hatást egyes csoportokra globális szinten?</p>	<p>1. Milyen mértékben járul hozzá ez az innováció a jövő nemzedékei számára egy igazságosabb és befogadóbb világ megteremtéséhez?</p> <p>2. Mennyire zárja ki ez az innováció annak lehetőségét, hogy olyan rendszerszintű előítéleteket rögzítsen, amelyek negatívan befolyásolják a jövő generációk lehetőségeit?</p>

<p>(D) Elszámoltathatóság és felelősség</p>	<p>1. Mennyire képes ez az innováció hosszú távon fenntartani a stabil és átlátható felelősségmegosztást, alkalmazkodva a változó körülményekhez?</p> <p>2. Mennyire biztosítja ez az innováció az elszámoltathatósági mechanizmusok rugalmasságát a jövőbeli kihívásokhoz esetén?</p>	<p>1. Milyen mértékben alakít ki ez az innováció olyan jogi és szabályozási keretrendszert, amely egyértelműen meghatározza a fejlesztők és alkalmazók felelősségét és elszámoltathatóságát?</p> <p>2. Mennyire támogatja és ösztönzi ez az innováció a felelősségteljes magatartást és döntéshozatalt, hozzájárulva ezzel egy erős felelősségi kultúra kialakulásához az érintettek körében?</p>	<p>1. Milyen mértékben tiszteletben tartja ez az innováció a különböző nemzetek és kultúrák szuverenitását és önrendelkezését?</p> <p>2. Mennyire biztosított az innováció hatásainak globális elszámoltathatósága?</p>	<p>1. Milyen mértékben előzi meg ez az innováció olyan elszámoltathatósági hiányosságok kialakulását, amelyek sebezhetővé vagy védtelenné tehetik a jövő generációit a nem szándékolt következményekkel szemben?</p> <p>2. Milyen mértékben biztosít ez az innováció olyan mechanizmusokat és keretrendszereket, amelyek lehetővé teszik a jövő generációi számára, hogy számonkérhetővé tegyék a fejlesztőket az innováció hosszú távú hatásaiért?</p>
<p>(E) Társadalmi és környezeti hatás</p>	<p>1. Milyen mértékben zárja ki ez az innováció annak a lehetőségét, hogy jövőben negatív társadalmi és környezeti következményei alakuljanak ki?</p> <p>2. Mennyire biztosítja ez az innováció, hogy hosszú távon ne vezessen az erőforrások kimerüléséhez?</p>	<p>1. Mennyire képes ez az innováció a társadalmi és környezeti kockázatok felmérésére és kezelésére?</p> <p>2. Mennyire biztosítja ez az innováció a fenntartható fejlődéshez való hozzájárulást?</p>	<p>1. Milyen mértékben van tekintettel ez az innováció a különböző régiók eltérő társadalmi és ökológiai rendszereire?</p> <p>2. Mennyire zárja ki ez az innováció annak lehetőségét, hogy a – akár nem szándékolt - hatásai legyenek a globális társadalmi vagy környezeti rendszerekre?</p>	<p>1. Milyen mértékben járul hozzá ez az innováció egy fenntartható világ megteremtéséhez a jövő generációi számára?</p> <p>2. Mennyire zárja ki ez az innováció annak lehetőségét, hogy korlátozza a jövő generációk számára elérhető erőforrásokat vagy károsítsa a környezetet?</p>
<p>(F) Kulturális integritás és identitás</p>	<p>1. Milyen mértékben képes ez az innováció hozzájárulni a kulturális örökség és a hagyományok megőrzéséhez a változó globális környezetben?</p> <p>2. Mennyire zárható ki annak a lehetősége, hogy jelen innováció a hatásai révén hosszú távon kulturális homogenizációhoz vagy a kulturális sokszínűség csökkenéséhez vezessen?</p>	<p>1. Milyen mértékben rendelkezik ez az innováció tervezett intézkedésekkel a kulturális örökség védelme és megőrzése érdekében?</p> <p>2. Mennyire biztosítja ez az innováció a kultúrák integritásának tiszteletben tartását és megőrzését?</p>	<p>1. Mennyire tartja tiszteletben és értékeli ez az innováció a világ különböző népeinek kulturális sokszínűségét?</p> <p>2. Mennyire zárja ki ez az innováció annak a lehetőségét, hogy hatásai aláássák a kulturális pluralizmus globális paradigmáját?</p>	<p>1. Milyen mértékben mozdítja elő ez az innováció a kulturális örökség megőrzését és továbbadását a jövő generációi számára?</p> <p>2. Mennyire zárja ki ez az innováció annak lehetőségét, hogy olyan helyzeteket hozzon létre, amelyek korlátozzák vagy beszűkítik majd a jövő generációk kulturális sokszínűségét, korlátozzák szabad identitásválasztásukat?</p>

A strukturált konszenzusereső megbeszélések során az értékelők célja az, hogy egy közös álláspontot alakítsanak ki az innováció etikai dimenzióinak értékelésében, tekintetbe véve a különböző nézőpontokat és érveket, véglegesítve az innováció etikai dimenzióinak pontszámait és a hozzájuk fűzött indoklásokat – és ezek alapján határozzák meg a végső responzibilitási indexet.

Az interszjektív értékelési folyamat csökkenti az egyéni szubjektivitásból eredő torzításokat, és elősegíti a több szempontú és kiegyensúlyozottabb kép kialakítását a vizsgált tárgyról. A felelős döntéshozatali mátrix így nem csupán az egyéni értékelések statisztikai összesítéséből áll össze, hanem egy komplexebb kvalitatív elemzési szempontokat is beépítő folyamat eredményeit használja fel a végső pontszámok meghatározásához.

A mátrix minden kérdéséhez kapcsolódóan, az értékelést végzők egy 1-5-ig terjedő skálán osztályozhatják a fejlesztésüket:

- (1) egyáltalán nem
- (2) inkább nem
- (3) semleges
- (4) inkább igen
- (5) teljes mértékben

A magasabb pontszám azt jelzi, hogy az adott etikai dimenzióhoz tartozó felelősségvállalási elvhez erősebben igazodik az adott innováció, az alacsonyabb pontszám egy kisebb igazodást jelöl.

Az egyes etikai dimenziókra vonatkozó általános pontszám kiszámításához a következő lépések követésével jutunk el:

- a) Az innováció értékelése minden egyes irányadó kérdésre vonatkozóan (pl. A-F sorokban lévő összes kérdés megválaszolása pontszámmal (az adható pontszám: 1-5));
- b) Az egyes cellákban szereplő kérdésekre adott értékelések összeadása (pl. A-III-as cellában lévő 1. és 2. kérdésre adott pontok összegzése $\rightarrow a_1+a_2=X$);
- c) A kapott összeg elosztása a kérdések számával, amely alapján megkapjuk az adott cella átlagos értékelését (pl. az A-III-as cellában a pontok összeértéke X , ezt osztjuk a kérdések számával(Y), jelen esetben 2db kérdés van, így 2-vel $\rightarrow X/Y=A_{III}$);
- d) Az adott sorra vonatkozó mind a négy cella átlagos értékelésének összeadása (pl. $A_I+A_{II}+A_{III}+A_{IV}=A_A$);
- e) A kapott összeg elosztása 4-gyel (alapelvek száma), hogy megkapjuk az adott etikai dimenzió általános pontszámát (pl. $A_A/4=A$ dimenzió általános pontértéke).

Az egyes etikai dimenziókra vonatkozó pontszámok segíthetnek a fejlesztőknek azonosítani az erősségeket és gyengeségeket az innovációjuk felelősséghez kapcsolódó etikai alapelveinek értékelési folyamatában. Ennek bemutatására a következőkben egy kitalált példát hozok, így tegyük fel, hogy egy innováció a következő értékeléseket kapja az „Emberi élet és méltóság tisztelete” (A) etikai dimenzióban:

9. táblázat. Példa az A etikai dimenzió kiszámításához (saját szerkesztés)

Emberi élet és méltóság tisztelete (A) etikai dimenzió					
Alapelv	1. kérdés (a ₁)	2. kérdés (a ₂)	Összeg (X)	Kérdések száma (Y)	Átlag (A _I , A _{II} ,...)
(I.) Imprevizibilis temporális hatások elve	4	3	7	2	3,5
(II.) Prekaucionalitás elve	5	4	9	2	4,5
(III.) Globális hatás-elv	3	3	6	2	3
(IV.) Intergenerációs hatások elve	4	2	6	2	3

Jelen esetben az Emberi élet és méltóság tisztelete (A) dimenzió átlagos értékeinek összege $A_I + A_{II} + A_{III} + A_{IV} = A_A$, azaz $3,5 + 4,5 + 3 + 3 = 14$, ebből fog következni, hogy ezen **etikai dimenzió általános pontértéke: $14/4 = 3,5$, azaz (A) = 3,5**

Ezek alapján az összes etikai dimenzióra ki tudjuk számolni az adott dimenzió értékét. A végső értékelőtáblázatban az egyes dimenziókra kapott pontok összeadódnak (jelen esetben $A + B + C + D + E + F$) és ezek átlaga (jelen esetben ezt az értéket 6-tal osztjuk, mert 6 etikai dimenziót vázoltunk fel) adja meg a fejlesztés **reszponzibilitási indexét**.

Ennek szemléltetésére is bemutatok egy példát, ahol tegyük fel, hogy az etikai dimenziók általános pontértékét a fentebb említett módszerrel már kiszámoltuk, ezeket az értékeket a táblázat tartalmazza:

10. táblázat. Példa a reszponzibilitási index (Ri) kiszámításához (saját szerkesztés)

Etikai dimenziók	Etikai dimenziók általános pontértéke
(A) Emberi élet és méltóság tisztelete	3,5
(B) Magánélet és adatvédelem	4
(C) Diszkrimináció és előítéletmentesség	3,5
(D) Elszámoltathatóság és felelősség	2
(E) Társadalmi és környezeti hatás	4
(F) Kulturális integritás és identitás	3

Összeadjuk az összes etikai dimenzióhoz tartozó általános pontértékeit, amelyet elosztunk az etikai dimenziók számával, amely így megadja a végső reszponzibilitási indexet (R_i).

$$R_i = \frac{(A + B + C + D + E + F)}{6} = \frac{(3,5 + 4,5 + 3,5 + 2 + 4 + 3)}{6} = 3,33$$

Az interszubjektív értékelés során minden értékelő rendelkezni fog egy saját R_i értékkel, amelyet a már fentebb említett módon számolhatunk ki. Annak érdekében, hogy megkapjuk az átlagos **interszubjektív reszponzibilitási index (iR_i) értékét** az összes értékelő R_i értékét össze kell adni, majd el kell osztani az értékelést végzők számával. Ennek bemutatására az alábbi táblázat és képlet ad segítséget:

11. táblázat. Példa R_i pontszámok az iR_i érték kiszámításához (saját szerkesztés)

Értékelő megnevezése	Értékelő R_i általános pontértéke
R_{i1}	3,33
R_{i2}	2,85
R_{i3}	4,17
R_{i4}	2,0
R_{i5}	4,42

$$iR_i = \frac{R_{i1} + R_{i2} + R_{i3} + \dots + R_{iN}}{N} = \frac{3,33 + 2,85 + 4,17 + 2,0 + 4,42}{5} = 3,35$$

azaz **$iR_i=3,35$** .

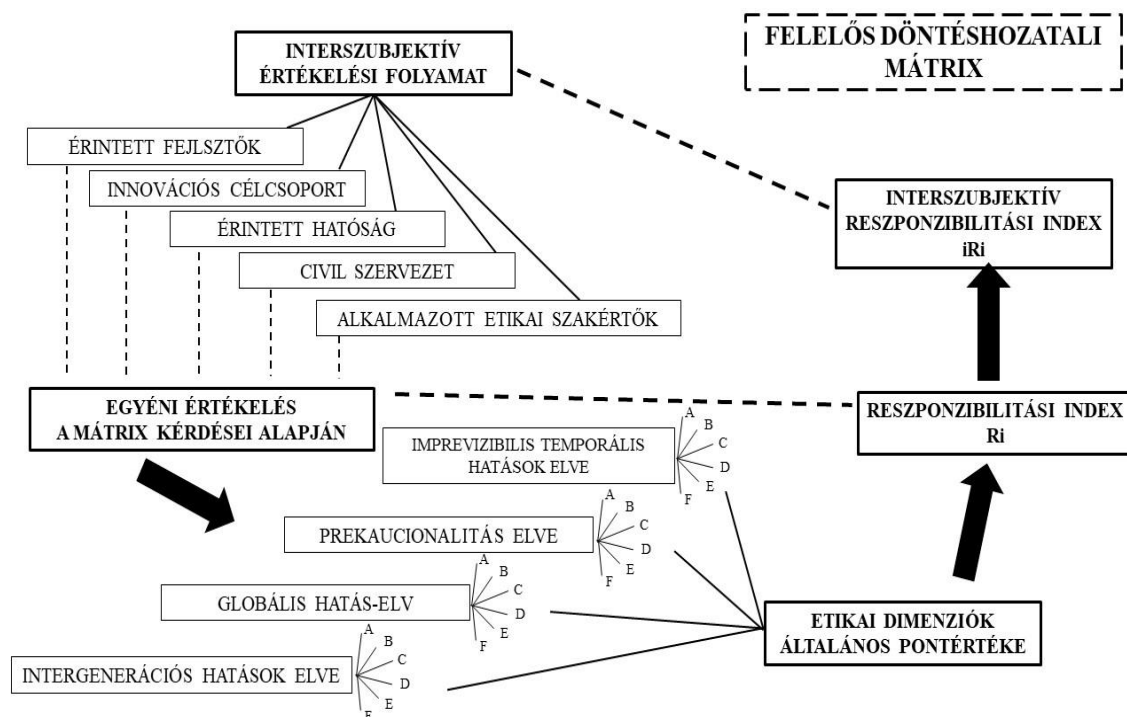
Ezt a számot a lenti skálán tudjuk értelmezni, ami az innovációk felelős etikai tervezésének értékelésére szolgál. A skála alapján öt kategóriába sorolhatjuk be az innovációkat, a nagyon alacsonytól a nagyon magas reszponzibilitási szintig, figyelembe véve az etikai dimenziók korábban fent már tárgyalt fenti spektrumát. Az értéktartományok határainak megadásánál figyelembe vettem a valószínűség számítás és hasonló skálarendszerek alapjait annak érdekében, hogy több értékelés esetén a torzító hatás ne érvényesüljön. Az értékeléshez az alábbi pontrendszert/értéktartományt határoztam meg:

12. táblázat. Az iRi értéktartományok határai (saját szerkesztés)

	iRi érték tartománya
Nagyon alacsony reszponzibilitás	1,0 – 1,9
Alacsony reszponzibilitás	2,0 – 2,4
Közepes reszponzibilitás	2,5 – 3,9
Magas reszponzibilitás	4,0 – 4,4
Nagyon magas reszponzibilitás	4,5 – 5,0

Az egyes kategóriák leírást nyújtanak a vizsgált fejlesztés, innováció etikai dimenzióiról, sajátosságairól és hiányosságairól, valamint az alkalmazásával járó potenciális kockázatokról. Az 1,0 és 1,9 közötti értékek egy nagyon alacsony etikai reszponzibilitásra utalnak, az adott innováció több dimenzióban jelentős hiányosságokat mutat és a fejlesztése/alkalmazása magas kockázatokkal járhat. A 2,0 és 2,4 közötti értékek szintén egy alacsony reszponzibilitási szintet jeleznek, az innováció itt is csak részben felel meg a felelős tervezés követelményeinek, és potenciálisan káros hatást eredményezhet több területen. Az ezt követő közepes reszponzibilitási szint (2,5 - 3,9) azt jelzi, hogy az innováció általánosságban megfelel az etikai elvárásoknak, de néhány területen további fejlesztésekre van szükség. A 4,0 és 4,4 közötti értékek már egy magas reszponzibilitást jelölnek: az innováció jól teljesít a legtöbb etikai dimenzióban és várhatóan a minimális kockázatot hordoz a vizsgált dimenziók tekintetében. Végül, a 4,5 és 5,0 közötti értékek kivételesen magas reszponzibilitási szintet mutatnak – ezeknél a számoknál az innováció kiemelkedően jól teljesít minden etikai szempont figyelembevételével.

Az értékelési rendszer alkalmazásával a fejlesztők árnyaltabb képet kaphatnak innovációjuk etikai profiljáról, és azonosíthatják a fejlesztendő területeket. Az index magasabb értéke azt jelzi, hogy az innováció erősebben igazodik a felelősségvállalás elveihez, míg alacsonyabb értéke hiányosságokat jelez ezen a téren.



12. ábra. A felelős döntéshozatali mátrix folyamatábrája (saját szerkesztés)

Jelen mátrix, gondolkísérletként egy alkalmazott etikai szemlélet jegyében született meg, és alkalmazásának korlátait főként az etikai kérdések kvantifikálhatóságának problémája adja. A komplex erkölcsi kérdések vizsgálata során a számszerűsítés kényszeréből adódó szimplifikáció oda vezethet, hogy az összetettebb etikai problémák rejtve maradnak. Mindezek mellett problematikus az értékelő személy szubjektív nézőpontja, ami szintén befolyásolhatja a végeredményt, és az értékelők közötti eltérések a mátrix konzisztenciájának megkérdőjelezését is jogosan felvetik – annak ellenére, hogy az interszubjektív értékelési rendszer ezen szubjektivitás fő problémáit egyértelműen kiküszöböli. Mindezeket kiegészíti a globális hatások előrejelzésének szinte lehetetlen feladata egy olyan új technológia fejlesztése esetében, mint a mesterséges intelligencia, ami magát a keretrendszert is folyamatos kihívás elé állítja, rendszeres felülvizsgálatra készítve az alkalmazókat. Mindezek ellenére a mátrixot jelen állapotában valid értékelési eszköznek tekinthetjük. A validitását két szempont adja: (a) a tartalmi validitás, amely az adott innováció kapcsán a lefedett etikai dimenziók és alapelvek átgondolt integrációját jelenti és a (b) fogalmi érvényesség, vagyis, hogy a kérdések és a vizsgált dimenziók összhangban vannak-e felelősség fogalmával.

Ami a fejlesztési lehetőségek kapcsán elmondható: [1] a mátrixban alkalmazott kérdések száma bővíthető a vizsgált terület komplexitásának függvényében; [2] minden

dimenzió kapcsán további finomhangolások és adaptációs felülvizsgálatok végezhetők; [3] a kritikus etikai szempontok rangsorolása segítheti a gyakorlati alkalmazást.

6.2.2 Adaptációs esettanulmány a felelős döntéshozatali mátrixhoz

A mátrix adaptációs esettanulmányában a mesterséges intelligencia egészségügyi innovációk területére fókuszáló alkalmazásán szeretném modellezni a felhasználás lehetőségeit, az orvosi diagnosztika területét választva a demonstrációra. Napjainkban számos fejlesztés folyik, amely az AI orvosi diagnosztikában történő felhasználásának a lehetőségeit kutatja. Az úgynevezett AI-powered Clinical Decision Support Systems (CDSSs - AI-vezérelt klinikai döntéstámogató rendszerek) amelyeket az egészségügyi szakembereket támogató technológiai rendszerekként értelmezhetünk, nagy pontosságú, bizonyítékalapú, betegszempontú diagnosztikai adatok szolgáltatása révén segítik az orvosi személyzetet a hatékony kezelés támogatásában (Elhaddad & Hamam, 2024). Az AI rohamos fejlődésével, egyre kevésbé kétséges, hogy képes lesz javítani és gyorsítani ezeknek a rendszereknek a pontosságát, komplex orvosi adatok akár egyidejű elemzését is elvégezve. A különféle adatforrások – demográfia, kórtörténet, laboratóriumi eredmények, valós idejű életjelek – integrálása, jelentős áttörést hozhat a diagnosztikában, lehetővé téve egészségügyi problémák korai felismerését. Ahogy a korábbiakban már tárgyalt területeken, a problémák itt is az érzékeny személyes adatok védelmével, a hozzáféréssel, az adatok pontosságával és az ehhez kapcsolódó felelősségi körökkel vannak (Al-Antari, 2023). A kérdés jelentőségét jól mutatja, hogy a tudományos szakfolyóiratok világából kilépve Meesha Dogan a Forbes magazin 2024 májusi számában tárgyalja az AI, különösen a gépi tanulás jelentőségét a diagnosztika területén az AI's Transformative Power In Healthcare Diagnostics című cikkében. Dogan felhívja a figyelmet arra, hogy a mesterséges intelligencia alkalmazása számos etikai aggályt felvet, megemlítve elfogultság és a lehetséges téves diagnózisok kapcsán felmerülő felelősség kérdését, éppúgy, mint a kiberbiztonsági kockázatokat és a betegadatok védelmét (Dogan, 2024).

A mobil egészségügyi (mHealth) és digitális egészségügyi technológiák legújabb eredményeinek és az új fejlesztéseknek a bemutatására fókuszáló The Journal of mHealth 2024 augusztusában közölte Alexander Ryan egészségpiaci szakértő írását, aki arra hívta fel a figyelmet, hogy a mesterséges intelligencia a prediktív analitika és a személyre szabott orvoslás terén is jelentős előrehaladást mutat. Az AI alapú prediktív analitika „képes azonosítani a krónikus betegségek kialakulásának kockázatával küzdő betegeket,

lehetővé téve a korai beavatkozást és a megelőző ellátást”, míg a személyre szabott orvoslásban az AI-algoritmusok „genetikai, környezeti és életmódbeli adatokat elemeznek, hogy a kezeléseket az egyes betegek egyedi jellemzőihez igazítsák, fokozva a kezelések hatékonyságát és minimalizálva a mellékhatásokat” (Ryan, 2024).

A fentiek tükrében az AI növekvő hatása az egészségügyi diagnosztikára tagadhatatlan, ahogy az ezzel járó etikai kérdések is várhatóan nagy számban kerülnek felszínre a következő években.

A következőkben egy hipotetikus, AI-alapú diagnosztikai fejlesztés kapcsán szemléltetem a mátrix működését. Első lépésként az interszubjektív értékelőcsoportot állítjuk össze, amelybe a fejlesztési folyamatban résztvevőkön túl, külső szakértőket is bevonunk: (a) orvosok és képző diagnosztikai szakemberek, (b) AI-fejlesztők és adattudósok, (c) bioetikai szakértők, (d) páciensek képviselői, illetve az (e) egészségügyi ellenőrző hatóság(ok) képviselői. Az értékelők előzetesen megkapják a fejlesztés összefoglalóit, amely tartalmazza az alkalmazási területekre, a hozzáférés módjaira, az adatfelhasználásra és adatkezelésre és a környezeti fenntarthatósági szempontokra vonatkozó komplex információkat. Minden értékelő egyénileg pontozza az innovációt a mátrixban szereplő összes etikai dimenzió mentén, majd közös megbeszélésen alakítják ki a végső pontszámokat.

A pontozólapok az egyéni megjegyzésekkel a következőképp néznek ki (a számok a mátrixban szereplő kérdések sorszámát jelölik és az adott lehetséges válaszokat):

13. táblázat. Az emberi élet és méltóság tisztelete- pontozás (saját szerkesztés)

	(A) Emberi élet és méltóság tisztelete	Pontszám (a _N)	Átlag (A _N)
(I.) Imprevizibilis temporális hatások elve	1. Az AI-alapú diagnosztika képes lehet időben észlelni súlyos betegségeket, ami pozitív hatással lehet az emberek életkilátásaira és életminőségére – ugyanakkor az emberi méltóság érzése csökkenhet a személyes kapcsolatok hiánya miatt	4	3,5
	2. A diagnosztikai hibák száma jelentős mértékben csökkenthető az AI bevonásával, de a technológia személytelensége negatív hatással lehet az egyéni méltóság érzésére	3	
(II.) Prekaucionalit ás elve	1. Az alkalmazott biztosítékok, mint például a személyes adatok védelmére vonatkozó intézkedések, hozzájárulhatnak az emberi méltóság megőrzéséhez	4	3,5
	2. Az AI diagnosztikai használata lehetőséget kínál egy megalapozottabb önálló döntéshozatalra, de a gépi döntéshozatali mechanizmusok növekvő szerepe hosszabb távon csökkentheti az egyéni autonómiát	3	
(III.) Globális hatás	1. Az AI rendszerek képesek lehetnek adaptálódni különböző kultúrák értékrendjéhez, de a globális szabványok elterjedése veszélyt jelenthet a kulturális sajátosságokra	3	3
	2. Az AI-alapú diagnosztikai rendszerek globális elterjedése elősegítheti az egészségügyi szolgáltatásokhoz való szélesebb körű hozzáférést, de fennáll a veszélye, hogy egyes csoportok így is kirekesztődnek	3	
(IV.) Intergenerációs hatások	1. Az AI-vezérelt diagnosztikai eszközök hosszú távon hozzájárulhatnak a betegségek korai felismeréséhez, ami pozitív hatással lesz a jövő generációk életminőségére	4	3,5
	2. Ha nem kezelik megfelelően az etikai kérdéseket, az innováció negatív következményekkel járhat, amelyek hosszú távon csökkenthetik a jövő generációk döntéshozatali autonómiáját és lehetőségeit	3	
		(A) Általános pontérték: 3,38	

14. táblázat. A magánélet és adatvédelem - pontozás (saját szerkesztés)

	(B) Magánélet és adatvédelem	Pontszám (b _N)	Átlag (B _N)
(I.) Imprevizibilis temporális hatások elve	1. Az AI-alapú rendszerek adatvédelmi szempontból a maiaknál sokkal biztonságosabbak lehetnek, de nem küszöbölik kis esetleges kibertámadások veszélyét	3	3,5
	2. Annak ellenére, hogy AI alkalmazások fejlett adattitkosítást használnak, a folyamatos frissítések szükségessége és a fenntarthatóság állandó kihívást jelent	4	
(II.) Prekaucionalitás elve	1. Jelen fejlesztés várhatóan komplex adatvédelmi mechanizmusokkal rendelkezik, amelyek képesek lesznek majd az új adatvédelmi kihívások kezelésére is	4	3,5
	2. Az AI-t használó technológiák adatgyűjtési módszerei és kapacitása aggodalmat kelthet a személyes adatok feletti hosszú távú emberi ellenőrzés kapcsán	3	
(III.) Globális hatás	1. Az AI rendszerek globális elterjedése során kihívást jelenthet a különböző országok adatvédelmi előírásainak alkalmazása	3	2,5
	2. Az AI rendszerek használata következtében megtörténhet az adatvédelmi egyenlőtlenségek növekedésének veszélye, különösen a kevésbé fejlett régiókban	2	
(IV.) Intergenerációs hatások	1. Az AI rendszerek működésének átláthatatlansága miatt a következő generációk számára nehezebbé válhat a személyes adatok feletti ellenőrzés visszaszerzése, ami tartós negatív hatással lehet a magánéletükre	3	2,5
	2. Az AI rendszerek által létrehozott nagyméretű adathalmazok és ezek potenciális piacosítása hosszú távon veszélyezteti az egyének magánéletét, a nem megfelelő védettség esetén	2	
		(B) Általános pontérték: 3,0	

15. táblázat. Diszkrimináció és előítéletmentesség (saját szerkesztés)

	(C) Diszkrimináció és előítéletmentesség	Pontszám (c _N)	Átlag (C _N)
(I.) Imprevizibilis temporális hatások elve	1. Az AI segíthet csökkenteni a diagnosztikai tevékenységek során felmerülő rejtett emberi előítéleteket, javítva ezzel az esélyegyenlőséget	3	3,5
	2. A rejtett algoritmikus torzítások hosszú távon fenntarthatják vagy súlyosbíthatják a negatív megkülönböztetést	4	
(II.) Prekaucionalitás elve	1. Az AI-alapú diagnosztikai rendszerek alkalmazása során lehetőség nyílik a rejtett algoritmikus előítéletek feltárására és csökkentésére, növelve a rendszer méltányosságát	4	3,5
	2. Az AI-alapú rendszerek képesek lehetnek a diagnosztikai folyamatok során a társadalmi befogadás erősítésére	3	
(III.) Globális hatás	1. Az AI-alapú diagnosztikai rendszerek képesek alkalmazkodni a különböző kultúrák egészségügyi gyakorlataihoz, de globális elterjedésükkel, fennállhat a veszélye, hogy a helyi egészségügyi hagyományok és kezelési módszerek háttérbe szorulnak	3	3
	2. Az AI képes lehet az eltérő kulturális kontextusok figyelembevételére, de kihívásokat jelenthet a kulturális érzékenység fenntartása	3	
(IV.) Intergenerációs hatások	1. Az AI-alapú rendszerek – a humán beavatkozás minimalizálásával – képesek lehetnek csökkenteni előítéleteket, ami hosszú távon elősegítheti egy igazságosabb és befogadóbb globális társadalom kialakulását	4	3,5
	2. Ezek a rendszerek rejtett torzításokat hordozhatnak magukban, amelyek hosszú távon is fennmaradhatnak, és hátrányosan befolyásolhatják a jövő generációk esélyeit	3	
		(C) Általános pontérték: 3,38	

16. táblázat. Elszámoltathatóság és felelősség - pontozás (saját szerkesztés)

	(D) Elszámoltathatóság és felelősség	Pontszám (d _N)	Átlag (D _N)
(I.) Imprevizibilis temporális hatások elve	1. Az AI rendszerek döntéshozatali mechanizmusai átláthatóbbá válhatnak, de a végső felelősség kérdése hosszú távon problémát jelenthet	3	3,5
	2. Amennyiben a fejlesztésbe beépítenek felelősségi mechanizmusokat, ezek fenntartása és naprakészen tartása állandó kihívást jelent majd	4	
(II.) Prekaucionalitás elve	1. Az AI-alapú technológiai innovációk bevezetése során szükséges a megfelelő jogi és szabályozási keretrendszer kialakítása. Szükséges meghatározni a felelősség és az elszámoltathatóság szempontjait.	4	3,5
	2. Az AI-alapú rendszerek felgyorsíthatják és leegyszerűsíthetik döntéshozatali folyamatokat, de az emberi tényező kihagyása hosszú távon veszélyeztetheti a felelősségi kultúra fenntartását	3	
(III.) Globális hatás	1. Problémát jelenthet az elszámoltathatóság, ha a technológia több országban (nem mindig azonos feltételek mellett) kerül alkalmazásra	3	3
	2. Hiányoznak a kötelező érvényű globális szabályozási keretek, ami jelen innováció esetében is egyes régiókban megnehezítheti az elszámoltathatóságot	3	
(IV.) Intergenerációs hatások	1. Fontos szempont, hogy a következő generációk rendelkezzenek olyan jogokkal és eszközökkel, hogy a technológia bármely később feltárt hibáját vagy nem szándékolt következményét korrigálni tudják	3	3
	2. Fontos, hogy rendelkezésre álljanak olyan mechanizmusok, amelyek biztosítják a fejlesztők elszámoltathatóságát hosszabb időtávon	3	
		(D) Általános pontérték: 3,25	

17. táblázat. Társadalmi és környezeti hatás - pontozás (saját szerkesztés)

	(E) Társadalmi és környezeti hatás	Pontszám (e _N)	Átlag (E _N)
(I.) Imprevizibilis temporális hatások elve	1. Amellett, hogy a fejlesztés javíthatja az egészségügyi rendszert, a technológia széleskörű elterjedése új társadalmi kihívásokat is hozhat (pl. hozzáférés)	3	2,5
	2. A fejlesztés magas energiaigénye és a technológiai hulladék növekedése jelentős környezeti terhelést okozhat	2	
(II.) Prekaucionalitás elve	1. A fejlesztés a beépített ellenőrző mechanizmusai révén lehetőséget nyújt a társadalmi és környezeti kockázatok pontosabb felmérésére, de a hosszú távú hatásai nem teljesen ismertek	3	2,5
	2. A fejlesztés energiaigénye és a kapcsolódó technológiai eszközök gyártásának környezeti hatása veszélyeztetheti a fenntarthatóságot	2	
(III.) Globális hatás	1. Figyelembe kell venni az adott fejlesztés alkalmazása során a különböző régiók társadalmi és ökológiai sajátosságait – egyes technológiák nem ugyanazon feltételekkel alkalmazhatók minden régióban	3	2,5
	2. Az alkalmazás során számolni kell a nem szándékolt hatásokkal, amelyek negatívan befolyásolhatják a globális társadalmi és környezeti rendszereket	2	
(IV.) Intergenerációs hatások	1. A fejlesztés hozzájárulhat a fenntarthatósághoz a pontosabb diagnosztika, a betegségmegelőzés és a szélesebb körű hozzáférés révén, de figyelmet kell fordítani az energiaigényre és az ökológiai lábnyomra is	3	2,5
	2. A fejlesztés jövőbeli széleskörű alkalmazása növelheti az energiafogyasztást és a technológiai hulladékot, ami veszélyeztetheti a jövő generációk erőforrásait	2	
		(E) Általános pontérték: 2,5	

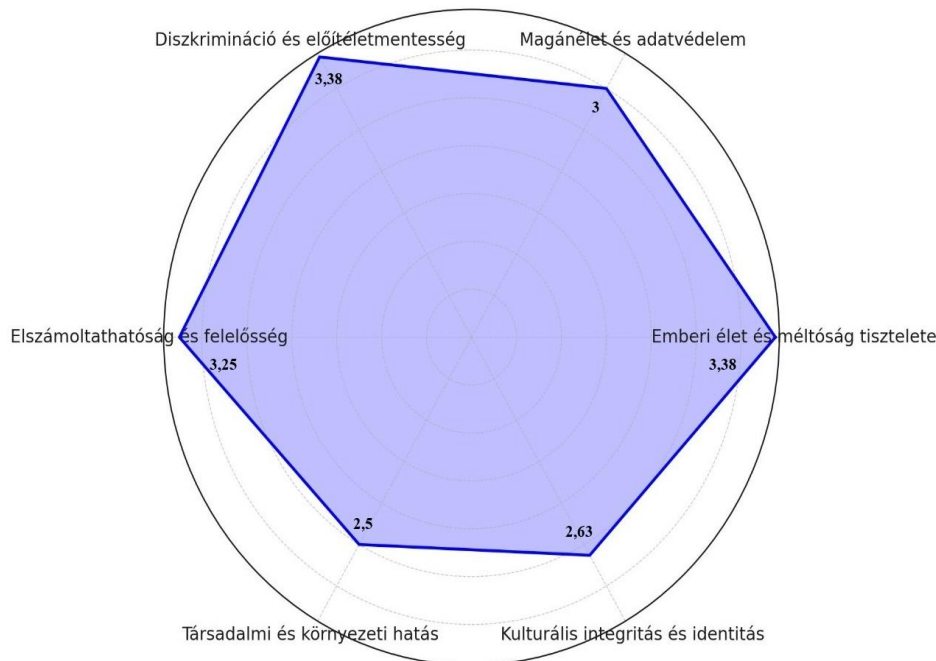
18. táblázat. Kulturális integritás és identitás (saját szerkesztés)

	(F) Kulturális integritás és identitás	Pontszám (f _N)	Átlag (F _N)
(I.) Imprevizibilis temporális hatások elve	1. A fejlesztés adaptálható lesz különböző kulturális kontextusokhoz, de fennáll a technológia homogenizációja miatt helyi módszerek és eljárások kerülnek veszélybe	3	2,5
	2. Az AI-alapú diagnosztika uniformizálhatja az orvosi gyakorlatot, ami csökkentheti a kulturális sokszínűséget és veszélyeztetheti pl. a hagyományos orvoslási módszereket	2	
(II.) Prekaucionalitás elve	1. Az AI-alapú diagnosztika beépítheti a működésébe a kulturális érzékenységet, de a globális sztenderdizáció veszélye így is fenyeget	3	3
	2. Az új technológiák bevezetése révén csökkenhet az orvoslás helyi sajátosságainak szerepe, mivel az egységesülő rendszerek kevésbé ösztönözik a regionális gyakorlatok és a hagyományos módszerek továbbfejlesztését – ez hosszabb távon a helyi innovációs potenciálra is kihathat	3	
(III.) Globális hatás	1. A fejlesztés képes lehet tiszteletben tartani a kulturális sokszínűséget, de a globális alkalmazás során nem lesz képes kezelni minden helyi specifikumot	3	2,5
	2. Az AI technológia alkalmazása során fennállhat a kulturális pluralizmus csökkenésének veszélye, különösen, ha a technológia nem veszi figyelembe a helyi kulturális sajátosságokat	2	
(IV.) Intergenerációs hatások	1. Az AI alapú rendszerek segíthetnek a kulturális örökség megőrzésében, ha megfelelően figyelembe veszik a helyi sajátosságokat és kulturális értékeket	3	2,5
	2. A fejlesztés hosszútávú alkalmazása, egy módszertani uniformizációt eredményezhet, ami hosszú távon munkaerőpiaci átalakulásokhoz, szakmák megszűnéséhez vezethet	2	
		(F) Általános pontérték: 2,63	

A táblázat – értékelő pontszámai alapján:

$$R_i = \frac{(A + B + C + D + E + F)}{6} = \frac{(3,38 + 3,0 + 3,38 + 3,25 + 2,5 + 2,63)}{6} = 3,02$$

azaz **R_i=3,02**.



13. ábra. A példa reszponzibilitási indexének radardiagramja (saját szerkesztés)

A fentiek alapján kalkulált érték egy közepes reszponzibilitási kategóriába esik, ami azt jelenti, hogy az értékelő véleménye alapján a tervezett fejlesztés általánosságban megfelel a felelősségetikai elvárásoknak, de néhány területen finomhangolásra van szükség. A fejlesztés etikai dimenzióinak átfogó értékelése mutatja, hogy bizonyos területeken átlagos vagy az átlagosnál erősebb teljesítményt nyújt, míg másokon kihívásokkal szembesül – ezek vizualizálására egy példa a különböző értékek radardiagramon való bemutatása. A radardiagram alapján jól látható, hogy a fejlesztés legnagyobb erősségei a diszkrimináció és előítéletmentesség, valamint az emberi élet és méltóság tisztelete terén mutatkoznak meg, azonban a társadalmi és környezeti hatások, illetve a kulturális integritás és identitás területein jelentős korrekciókra van szükség. Ezek az alacsonyabb pontszámok utalhatnak arra, hogy a fejlesztés nem veszi kellően figyelembe a helyi társadalmi és kulturális sajátosságokat, illetve a környezeti fenntarthatóságot, ami hosszú távon veszélyeztetheti a társadalmi elfogadottságát és a jogi szabályozásokkal is konfliktusba kerülhet. Az elszámoltathatóság és felelősség terén

is viszonylag jól teljesít, jelezve, hogy megfelelő eszközök biztosítják a döntéshozatal átláthatóságát és a felelősség megosztását.

A fentihez hasonló vizuális értékelési eszközök könnyű áttekinthetőségük révén gyorsítják az értékelés záró szakaszát, amikor egy strukturált konszenzuseresési folyamat indul, amely során az interszubjektív csoport tagjai megvitatják a pontszámokat és indoklásokat. A cél az, hogy a különböző nézőpontok figyelembevételével közös álláspontot alakítsanak ki a végső pontszámokról vagy szükség esetén a pontszámok átlagait használják a végleges érték meghatározásához (ennek az iR_i -értéknek a kiszámítását az előző szakaszban részleteztem). Az evaluációs folyamat lezárultával, a konszenzuálisan meghatározott végső pontszámokat és azok részletes indoklását egy átfogó jelentés összegzi, amely tartalmazza az egyes etikai dimenziók kapcsán felvetődött kulcsfontosságú észrevételek szöveges kifejtését, a pontszámok közötti potenciális diszkrpanciák okainak vizsgálatát, valamint a kialakított közös álláspontot és a fejlesztés végső reszponzibilitási indexét (iR_i).

6.3 Összefoglalás

A reszponzibilitási index célja, hogy áttekintést nyújtson az adott innováció felelősségetikai szempontú megfelelőségéről, feltárja a hiányosságokat és a záró jelentésben javaslatokat fogalmazzon meg a lehetséges fejlesztési irányokra és a felmerült kihívások megoldására vonatkozóan.

A fejezet zárásaként hangsúlyozni szeretném, hogy a felelős döntéshozatali mátrix alkalmazása nem kizárólag egy fejlesztés felelősségetikai megfelelőségének vizsgálatára korlátozódik, hanem az innovációs folyamat során érintett összes szereplő perspektíváját képes megjeleníteni, ezáltal egy komplex képet kialakítani. A mátrix alkalmazásával képesek vagyunk sikeresen azonosítani mindazon dimenziókat, amelyekben egy innováció jól teljesít, valamint azokat, ahol további figyelemre és beavatkozásra van szükség a hatások kezelése érdekében. A mátrix implementálása révén egy olyan átfogó etikai keretrendszer jön létre, amelyben a felelősség szempontjai kapnak prioritást, egy többdimenziós értékelési rendszerben határozva meg azokat a princípiumokat és cselekvési modelleket, amelyek összhangban vannak a szabályozó dokumentumokban és ajánlásokban részletesen kifejtett emberi méltóság, társadalmi igazságosság és környezeti fenntarthatóság alapelveivel – úgy, hogy explicit módon támogatja a felelős innovációt és a fenntartható technológiai fejlődés alapelvét.

VII. Összegzés és konklúzió

A doktori értekezésem célja az volt, hogy átfogó képet adjak az innováció és a felelősség fogalmának komplex kapcsolatáról, valamint megtaláljam azokat a filozófiai és etikai kereteket, amelyek segítséget nyújthatnak a felelős innováció gyakorlatának kialakításához. Egy hangsúlyosan interdiszciplináris megközelítést alkalmazva törekedtem arra, hogy mélyreható elemzését kínáljam a választott témának.

Az első kutatási kérdésem az innováció és a felelősség fogalmának definíciós problémáira és azok kortárs filozófiai diskurzusban elfoglalt helyére irányult. *(K₁) Milyen definíciós problémák merülnek fel az innováció és felelősség kontextusában, és hol van ezeknek a fogalmaknak a helye a kortárs filozófiai diskurzusban?* Igyekeztem rávilágítani arra, hogy mind az innováció, mind a felelősség fogalma többretegű és kontextusfüggő, ami megnehezíti az egységes definíciók kialakítását: az innováció fogalma a technológiai és gazdasági megközelítésektől a filozófiai értelmezésekig terjed, míg a felelősség fogalma magában foglalja az egyéni elszámoltathatóságtól a kollektív és intergenerációs dimenzióig terjedő aspektusokat egyaránt. Amellett érveltem, hogy a fogalmak komplexitásának elismerése elengedhetetlen a felelős innováció kereteinek kialakításához, és javaslatot tettem az innováció és a felelősség átfogó, multidimenzionális értelmezésére, saját definíciók formájában.

A második kutatási kérdésem az innováció fogalmának történeti fejlődésére és a kortárs etikai problémák értelmezéséhez való hozzájárulására összpontosított. *(K₂) Miként járul hozzá az innováció fogalom történeti fejlődésének vizsgálata az innováció által generált kortárs etikai problémák tárgyalásához?* A dolgozatban egy filozófiatörténeti áttekintést nyújtottam az innováció értelmezésének változásairól, az ókori görög filozófiától a modern korig. Az elemzés során feltártam, hogy az innováció megítélése az idők során az ambivalenciától a pozitív értelmezésig terjedt, és mindig szorosan összefonódott az adott korszak társadalmi-kulturális kontextusával. A történeti perspektíva rávilágított arra, hogy az innováció etikai vonatkozásainak vizsgálata nem újkeletű, az újítás kérdése implicit módon a kezdetek óta jelen van a nyugati filozófiai gondolkodásban – ez a felismerés hozzájárul a kortárs etikai dilemmák elmélyültebb megértéséhez és a történetileg kialakult fogalmi keretek továbbfejlesztéséhez.

A harmadik kutatási kérdés a felelősségetika és az innovációetika helyét és súlyát vizsgálta a kortárs tudományos diskurzusban. *(K₃) Milyen helyet foglal el és milyen súllyal jelenik meg a felelősségetika és az innovációetika a kortárs tudományos diskurzusban?* A dolgozatban elvégeztem egy bibliometriai elemzést, amely segítségével kimutattam az egyes témák (felelősség, innováció, etika) iránti növekvő érdeklődést, különösen az

orvostudományok, a társadalomtudományok és az informatika területén. Ugyanakkor a kutatás rámutatott arra is, hogy az innovációetika és a felelősségetika mint önálló diszciplína még nem teljesen kiforrott, és további elméleti és módszertani tisztázást igénylő kérdések felmerülnek. Abból kiindulva, hogy mindkét etika inherens módon interdiszciplináris terület, vizsgálni kell, hogy járulhat hozzá a hibridizáció a fejlődésükhöz, milyen módszertani kihívásokkal kell szembenézni az ilyen megközelítések alkalmazása során. Fontos feladat az innovációetika és a felelősségetika viszonyát tisztázni más ágazati etikákkal (bioetika, környezeti etika vagy mérnöki etika), az egyes területek pontos lehatárolása, az önálló elméleti és módszertani sajátosságok elkülönítése miatt.

A negyedik kutatási kérdés azokra a filozófiai és etikai elméletekre összpontosított, amelyek kulcsszerepet játszanak a felelős innováció gyakorlati kereteinek kialakításában. *(K₄) Milyen filozófiai és etikai elméletek játszanak kulcsszerepet a felelős innováció, mint praxis kereteinek kialakításában?* A dolgozatban számos koncepciót áttekintettem, illetve részletesebben tárgyaltam Hans Jonas felelősségetikáját és a félelem heurisztikájának elméletét, mint a felelős döntéshozatal egy lehetséges alapját, kiemelve a jövő generációk iránt érzett felelősség, az elővigyázatossági elv és a fenntarthatóság szerepét a felelős innováció normatív kereteinek meghatározásában. Amellett érveltem, hogy ezen elméletek szintézise és egy holisztikus megközelítés elengedhetetlen ahhoz, hogy a felelős innováció gyakorlati alkalmazásának kereteit kialakítsuk. A felelős innovációnak a gyakorlatában reflektálnia kell más filozófiai és etikai elméletre is (konzekvencializmus, deontológia, diskurzusetika), és kontextusérzékeny módon kell egyesítenie a különböző megfontolásokat.

Az ötödik kutatási kérdés egy felelősségetikai modell kialakításához szükséges etikai alapelvekre és módszertani lehetőségekre irányult. *(K₅) Melyek azok az etikai princípiumok és követelmények, amelyek elengedhetetlenek egy konzisztens felelősségetikai modell kialakításához az innovációs folyamatok etikai szempontú megfelelésének értékelésére? Milyen módszertani lehetőségek és korlátok merülnek fel egy ilyen modell létrehozása során?* Erre vonatkozóan a dolgozatban javaslatot tettem egy lehetséges mérőeszköz, az általam felelős döntéshozatali mátrixnak nevezett modell bevezetésére, amely a hosszú távú következmények, az elővigyázatosság, a globális hatások és az intergenerációs felelősség dimenzióit integrálja. A mátrix célja, hogy strukturált keretet nyújtson az innovációs folyamatok etikai megfelelésének értékeléséhez, és elősegítse a felelős döntéshozatalt. A dolgozatban hangsúlyoztam a mátrix alkalmazásának korlátait, mint például az etikai kérdések kvantifikálásának nehézségeit és a szubjektivitás problémáját, ugyanakkor rámutattam a módszer előnyeire is, mint például az átfogó és többszemponú értékelés lehetőségére.

(K₆) Milyen etikai dilemmákat és társadalmi kihívásokat generál a mesterséges intelligencia fejlődése és széleskörű alkalmazása? Hogyan képes a felelős innováció hozzájárulni a mesterséges intelligencia fejlesztésének és alkalmazásának etikai szempontú szabályozásához? A hatodik kutatási kérdés a mesterséges intelligencia fejlődéséből eredő etikai dilemmákra és társadalmi kihívásokra, valamint a felelős innováció szerepére fókuszált, amire választ keresve a dolgozatban részletesen elemeztem a mesterséges intelligencia alkalmazásának etikai vonatkozásait a biztonságpolitika, az oktatás és a munka világában, rámutatva olyan kulcsfontosságú kérdésekre, mint az adatvédelem, a diszkrimináció és az emberi kontroll és ágencia elvesztése. Hangsúlyoztam a felelős innováció jelentőségét a mesterséges intelligencia fejlesztésének és alkalmazásának etikai szabályozásában, és javaslatot tettem egy új értékelési keretrendszer bevezetésére a mesterséges intelligenciát használó fejlesztések esetében, amit egy konkrét esettanulmányban mutattam be egy AI-alapú diagnosztikai innováción.

A bemutatott eredmények, érvek és esettanulmányok nemcsak a kutatási kérdésekre adnak választ, hanem egyben hozzájárulnak a disszertáció elején felállított két központi hipotézis vizsgálatához is. Az *első hipotézis* szerint az etikai elvek következetes alkalmazását az innovációs folyamatokban akadályozza egy egységes etikai elméleti keret hiánya. A dolgozat eredményei alátámasztják ezt a hipotézist, rámutatva a különböző filozófiai és etikai megközelítések sokféleségére és azok eltérő gyakorlati implikációira. Ugyanakkor a dolgozatban amellet érvelek, hogy a különböző perspektívák integrációja és egy holisztikus felelősségetikai keret kialakítása hozzájárulhat a koherencia és a következetesség növeléséhez.

A *második hipotézis* szerint a felelősségetika alapelvei integrálhatók egy értékelési modellbe, amely képes mérni az innovációs folyamatok etikai megfelelőségét és elősegíteni a felelős innováció gyakorlatát. A dolgozat igazolja ezt a hipotézist a felelős döntéshozatali mátrix bemutatásával, amely a felelősségetika dimenzióit operacionalizálja és gyakorlati alkalmazhatóságát demonstrálja az AI-alapú innovációk példáján.

A célom az volt, hogy a kutatásommal hozzájáruljak az innováció és a felelősség fogalmának megértéséhez, feltárva azok történeti, kulturális, társadalmi és etikai dimenzióit. Próbáltam rávilágítani a felelősségetika és az innovációetika – viszonylagos marginalitása mellett – növekvő jelentőségére a kortárs tudományosságban, és javaslatot tenni egy átfogó felelősségetikai keretre, amely elősegítheti a felelős innováció gyakorlatának elterjedését. A dolgozat eredményei alapul szolgálhatnak további kutatásokhoz az innováció etikai vonatkozásai kapcsán, valamint hozzájárulhatnak a felelős innováció elveinek és gyakorlatának fejlesztéséhez a különböző kutatási területeken. A jövőben tervezem a felelősségetikai keret

finomítását, foglalkozni akarok a mátrix tesztelésével más innovációs kontextusokban, valamint a felelős innováció szervezeti és szabályozási aspektusaival.

Ahogy a dolgozat egészében láthattuk, az új innovációk alkalmazása során változatos morális dilemmák merülnek fel (autonómia vs. biztonság, hatékonyság vs. méltányosság, rövid távú előnyök vs. hosszú távú kockázatok, egyéni jogok vs. társadalmi jólét), amelyek megkövetelik az etikai szempontok erőteljesebb integrálását a tudománypolitikába, amelynek rugalmas alkalmazkodása az új kihívásokhoz elengedhetetlen. A technológiai fejlődés globális természete megköveteli a nemzetközi együttműködést az etikai normák kialakításában: az emberi jogok, a szabadságjogok és az általános emberi biztonsági igények közötti egyensúly megtalálása kulcsfontosságú az etikus innovációhoz, itt gyakorlati etikai irányelvek kidolgozása segíthet az innovációk biztonságos alkalmazásának tervezésében és irányításában. A globális hatalmi központokban és az innovációs centrumokban ma meghozott döntések helyi szinteken is hosszútávú következményekkel járnak, hatással lesznek az egyéni szabadságjogokra, a társadalmi struktúrákra és a nemzetközi kapcsolatokra egyaránt. Ezért, ahogyan az egyre inkább digitalizálódó jövő felé haladunk, elengedhetetlen, hogy komplex, sokszereplős párbeszédet folytassunk, és olyan kereteket dolgozzunk ki, amelyek az innovációt és az etikai integritást egyaránt előtérbe helyezik. Egy olyan egyensúly megteremtésére kell törekednünk, amely elismeri az innováció gazdaság és társadalomformáló erejét, ugyanakkor fenntartja közös emberi értékeink és biztonsági szükségleteink elsőbbségét.

Az innováció és a biztonság összeegyeztetésének dinamikus folyamatában a politikai döntéshozó fórumoknak kiemelt szerep jut majd a jövőben. Miközben további technológiai ugrások előtt állunk, az etikai megfontolásaink és a holisztikus biztonság iránti elkötelezettségünk fogja alakítani a modern tudománypolitika közegét.

A kortárs tudományos gyakorlat etikai dilemmáit elemezve korábban rámutattam (Szilágyi, 2024a, pp. 147-148), hogy a modern tudományos közösség jelentős ellentmondásokkal néz szembe, mivel egyre nagyobb hangsúlyt kap a felelősség és a hosszú távú tervezés, miközben továbbra is az innováció és a felfedezés iránti erős késztetés hajtja. A korlátozások nélküli innováció, bár jelentős társadalmi előnyökkel jár, etikai megfontolások hiányában veszélyeket is hordoz magában. E veszélyek kezelése érdekében szükséges, hogy a tudományos közösség – a tudományos testületek, a kutatások finanszírozását nyújtó cégek, a tudománypolitikát befolyásoló kormányzati szervek és egyes vezető tudományos szaktekintélyek – részletes egyeztetéseket folytassanak és hatékonyan együttműködjenek.

Annak érdekében, hogy a tudomány jövőbeli irányát a nagyobb felelősségvállalás felé tereljük, elengedhetetlen, hogy további vitákat, és kutatásokat folytassunk ezekről a kritikus

kérdésekről, és ezekbe a vizsgálatokba több tudományág szakembereit be kell vonni. Jelenleg a nyitott és etikus tudomány eszményei, amelyek a társadalom számára hasznot hoznak, folyamatosan fenyegetettségben vannak. Az etikát és a fenntarthatóságot előtérbe helyező tudományos kultúra előmozdítása érdekében alapvető fontosságú új normák és útmutatások meghatározása, valamint olyan képzések biztosítása, amelyek ösztönzik a társadalmi felelősségvállalást. E nehézségek megoldása érdekében az innovációs kényszert fel kell váltania egy együttműködést és fenntartható fejlődést fókuszába állító új paradigmának.

Köszönetnyilvánítás

Mindenekelőtt szeretném kifejezni hálámat és köszönetemet Dr. Dudok Fanninak, aki nem csupán értő olvasóként, hanem lelki támaszként is mellettem állt a doktori dolgozat megírása során. Támogatása és bátorítása nélkülözhetetlen volt számomra.

Őszinte köszönettel tartozom témavezetőmnek, Prof. Dr. Nemeskéri Zsoltnak, akinek ötletei és lelkesedése végig inspirációt jelentettek. Szakmai iránymutatása és konstruktív visszajelzései nagyban hozzájárultak a disszertáció elkészüléséhez.

Hálás vagyok a Humántudományi Innovációs Kutatócsoport és a Filozófia Doktori Iskola vezetőjének, Prof. Dr. Szécsi Gábornak az innováció és etika területét érintő értékes tudományos útmutatásáért. Meglátásai és tanácsai segítettek a téma mélyebb megértésében és kidolgozásában. Hasonló hálával gondolok Prof. Dr. Boros Jánosra, aki a Doktori Iskola volt vezetőjeként sokat tett azért, hogy ösztöndíjas időszakom alatt egy befogadó tudományos közegben dolgozhassak.

Köszönöm Függ Zsolt Péternek, aki a Kooperatív Doktori Programban munkahelyi vezetőként gyakorlati támponokat adott munkámhoz. Az ő tapasztalatai és iránymutatása fontos szerepet játszottak a kutatás gyakorlati vonatkozásainak feltárásában.

Végül, de nem utolsósorban, szeretnék köszönetet mondani Prof. Dr. Máté-Tóth Andrásnak, akitől a tudományműveléshez szükséges ismereteim nagy részét kaptam. Az ő oktatói és kutatói példája és a tőle kapott módszertani iránymutatások meghatározók voltak számomra a munkám során.

Mindannyiuknak hálával tartozom a támogatásukért, tudásukért és az irántam tanúsított bizalmukért. Nélkülük ez a doktori dolgozat nem jöhetett volna létre.

Felhasznált irodalom

- Aaronson, S. (2013). *Quantum computing since Democritus*. Cambridge University Press.
- Adams, R., Bessant, J. & Phelps, R. (2006). Innovation management measurement: A review. *International Journal of Management Reviews*, 8(1), 22.
- Aguinis, H. & Glavas, A. (2012). What we know and don't know about corporate social responsibility: A review and research agenda. *Journal of Management*, 38(4), 932-968.
- Al-Farabi (1997). Az eszményi város lakosainak nézeteiről. In *Arab filozófia*. Szöveggyűjtemény II. Pázmány Péter Katolikus Egyetem Bölcsészettudományi Kar, Piliscsaba.
- Al-Ghazali (1993). *Ihya ulum al-din [The revival of religious sciences]*. Dar al-Khayr.
- Al-Ghazali (2000). *The Incoherence of the Philosophers (Tahafut al-Falasifah)*. Marmura, M. E. (fordító). Brigham Young University Press.
- Almond, B. (1996). „Applied Ethics”. In T. Mautner (Ed.) *Dictionary of Philosophy*, Penguin.
- Almond, B. (Ed.) (1995). *Introducing Applied Ethics*. Cambridge, Wiley-Blackwell.
- Aquinói Szent Tamás (2011). *A Summa Theologiae kérdései a jogról*. Szent István Társulat.
- Arisztotelész (1969). *Politika*. Gondolat Kiadó.
- Arisztotelész (1987). *Nikomakhoszi etika*. Európa könyvkiadó.
- Arntz, M., Gregory, T. & Zierahn, U. (2016). The risk of automation for jobs in OECD countries: A comparative analysis. *OECD Social, Employment and Migration Working Papers*, (189). OECD Publishing.
- Attar, S. (2007). *The Vital Roots of European Enlightenment: Ibn Tufayl's Influence on Modern Western Thought*. Lexington Books.
- Atzori, L., Iera, A. & Morabito, G. (2017). Understanding the Internet of Things: Definition, potentials, and societal role of a fast-evolving paradigm. *Ad Hoc Networks*, 56, 122-140.
- Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3-30.
- Babich, B. E. (1994). *Nietzsche's Philosophy of Science: Reflecting Science on the Ground of Art and Life*. State University of New York Press.
- Bacon, F. (1901). *The Advancement of Learning*. (Ed. Joseph Devey) P.F. Collier and Son.
- Bacon, F. (1987). *Esszék*. Julow V. (fordító). Európa Könyvkiadó.
- Bacon, F. (2001). *Novum Organum*. Csatlós, J. (fordító). Új Atlantisz. Sarkady, J. (fordító). Lazi Bt., Szeged.
- Bacon, F. (2019). Sir Francis Bacon: Valerius terminus: A természet magyarázata Hermes Stella megjegyzéseivel. *Magyar Filozófiai Szemle*, 63(3), 218–225.
- Bard, A. & Söderqvist, J. (2002). *Netocracy: The new power elite and life after capitalism*. Reuters.
- Barnett, H. G. (1942). Invention and Cultural Change. *American Anthropologist*, 44(1), 14–30.
- Barnett, H. G. (1953). *Innovation: The Basis of Cultural Change*. McGraw-Hill.
- Beauchamp, T. L. & Childress, J. F. (1979). *Principles of biomedical ethics*. Oxford University Press.
- Beauchamp, T. L. & Childress, J. F. (2012). *Principles of biomedical ethics* (7th ed.). Oxford University Press.

- Beauchamp, T. L. (2003). The Nature of Applied Ethics. In R. G. Frey & C. H. Wellman (Eds.), *A Companion to Applied Ethics* (pp.1-16). Wiley-Blackwell.
- Beauchamp, T. L. (2007). The 'four principles' approach to health care ethics. In R. E. Ashcroft (Ed.), *Principles of health care ethics*. Hoboken, Wiley.
- Bedeau, H. (2001). „Applied Ethics.” In L. C. Becker & C. B. Becker (Eds.), *Encyclopedia of Ethics* (pp. 80-81). Routledge.
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.
- Bentley, S. V., Naughtin, C. K., McGrath, M. J., Irons, J. L. & Cooper, P. S. (2024). The digital divide in action: How experiences of digital technology shape future relationships with artificial intelligence. *AI and Ethics*.
- Bessen, J. E. (2015). *Learning by doing: The real connection between innovation, wages, and wealth*. Yale University Press.
- Blok, V. (2021). Philosophical reflections on the concept of Innovation. In B. Godin, G. Gaglio & D. Vick (Eds.). *Handbook on Alternative Theories of Innovation* (pp. 354-367). Northampton, Verenigd Koninkrijk.
- Boddington, P. (2017). *Towards a code of ethics for artificial intelligence*. Springer.
- Bostrom, N. & Yudkowsky, E. (2014). The ethics of artificial intelligence. In *The Cambridge handbook of artificial intelligence* (pp. 316-334). Cambridge University Press.
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Boudreau, J. W. & Cascio, W. F. (2017). Human capital analytics: Why are we not there? *Journal of Organizational Effectiveness: People and Performance*, 4(2), 119-126.
- Bourban, M. & Rochel, J. (2020). Synergies in innovation: Lessons learnt from innovation ethics for responsible innovation. *Philosophy and Technology*, 34(2), 373-394.
- Brainard, J. (2018). Rethinking retractions. *Science*, 362(6413), 390-393.
- Brassier, R. (2007). *Nihil Unbound: Enlightenment and Extinction*. Palgrave Macmillan.
- Brennen, J. S. & Kreiss, D. (2016). Digitalization. In *The International Encyclopedia of Communication Theory and Philosophy* (pp. 1-11). Wiley-Blackwell.
- Brey, P. (2021). Understanding Engineering Design and Its Social, Political, and Moral Dimensions. In S. Vallor (Ed.), *The Oxford Handbook of Philosophy of Technology* (pp. 400-401). Oxford University Press.
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Anderson, H. (2018, Eds.). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. University of Oxford.
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., Nori, H., Panagi, H., Ribeiro, M.T. & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. *arXiv preprint arXiv:2303.12712*.
- Buolamwini, J. & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification..Proceedings of the 1st Conference on Fairness, Accountability and Transparency, in Proceedings of Machine Learning Research. 81:77-91. <https://proceedings.mlr.press/v81/buolamwini18a.html>.
- Burget, M., Bardone, E. & Pedaste, M. (2017). Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and Engineering Ethics*, 23(1), 1-19.

- Burnet, J. (1928). *Greek Philosophy – From Thales to Plato*. London, Macmillan and Co.
- Callahan, J. F. (1986). Richard Peter McKeon 1900-1985. *Journal of the History of Ideas*, 47(4), 653–662.
- Callicott, J. B. (1990). The Case against Moral Pluralism. *Environmental Ethics*, 12(2), 99-124.
- Carroll, A. B. (1991). The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders. *Business Horizons*, 34(4), 39-48.
- Cemil, A. (2022). Al-Ghazālī and the emergence of modern science. In *Science and Technology in the Islamic World*. Proceedings of the XXth International Congress of History of Science. (Liege, 20-26 July, 1997). Turnhout, Belgium Brepols Publishers.
- Chalmers, D. J. (2010). The singularity: A philosophical analysis. *Journal of Consciousness Studies*, 17(9-10), 9-10.
- Coeckelbergh, M. (2017). *Using Words and Things: Language and Philosophy of Technology*. Routledge.
- Cohen, H. F. (1994). *The Scientific Revolution. The Historiographical Inquiry*. Chicago, The University of Chicago Press Chicago
- Cooper, J. R. (1998). A multidimensional approach to the adoption of innovation, *Management Decision*, 36 (8), 493-502.
- Corbin, H. (1993). *History of Islamic Philosophy*. Kegan Paul International.
- Crawford, K. & Calo, R. (2016). There is a blind spot in AI research. *Nature*, 538(7625), 311-313.
- Csizmadia, T. (2023). *Innováció versus minőségmenedzsment - avagy az oximoron feloldása*. Budapest, Akadémiai Kiadó.
- Daugherty, P. R. & Wilson, H. J. (2018). *Human + Machine: Reimagining work in the age of AI*. Harvard Business Review Press.
- Davenport, T. H., Guha, A., Grewal, D. & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1), 24-42.
- de Boer, T. J. & Jones, E. R. (1933). *The history of philosophy in Islam*. London, Luzac & co.
- Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer Nature.
- Dogan, M. & Pahre, R. (1999). *Creative marginality: Innovation at the intersections of social sciences*. Westview Press.
- Dogan, M. (2003). A politikatudomány és a többi társadalomtudomány. In R. E. Goodin & K. Hans-Dieter (Eds.), *A politikatudomány új kézikönyve* (pp. 103–136). Osiris.
- Downs, G. W. & Mohr, L. B. (1976). Conceptual issues in the study of innovation. *Administrative Science Quarterly*, 21(4), 700–714.
- Drucker, P. (1999). Knowledge-worker productivity: The biggest challenge. *California Management Review*, 41(2), 79-94.
- Drucker, P. F. (1954). *The practice of management* (1st ed.). Harper & Row Publisher.
- Drucker, P. F. (1985). *Innovation and entrepreneurship*. Harper & Row Publishers.
- Duff, R. (1998). Responsibility. In *The Routledge Encyclopedia of Philosophy*. Taylor and Francis. (DOI: 10.4324/9780415249126-L085-1)
- Edelstein, L. (1967). *The Idea of Progress in Classical Antiquity*. Baltimore.

- Efrati, A. & Victor, J. (2023). OpenAI made an AI breakthrough before Altman firing, stoking excitement and concern. <https://www.theinformation.com/articles/openai-made-an-ai-breakthrough-before-altman-firing-stoking-excitement-and-concern>
- Elliott, K. C. (2017). *A tapestry of values: An introduction to values in science*. Oxford University Press.
- Ellul, J. (1964). *The Technological Society*. Vintage Books.
- Ellul, J. (1994). Az ember a technika rendszerében. In J. A. Tillmann (Ed.), *A későújkor józansága II. Olvasókönyv a tudományos-technikai világfelszámolás köréből* (pp. 134-150). Göncöl kiadó.
- Etzioni, A. (1988). *The Moral Dimension, Toward a New Economics*. The Free Press.
- Etzioni, A. (1993). *The Spirit of Community: Rights, Responsibilities and the Communitarian Agenda*. Crown.
- Ewing, A. C. (1953). *Ethics*. London, English Universities Press.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Schafer, B. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707.
- Frankena, W. K. (1973). *Ethics*. (2nd edition.) Prentice-Hall.
- Frey, C. B. & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254-280.
- Frison, G. (1988). Technical and technological innovation in Marx. *History and Technology*, 6(4), 299–324.
- Fuchs, C. (2014). *Social media: A critical introduction*. SAGE Publications.
- Garavan, T. N., Watson, S., Carbery, R. & O'Brien, F. (2016). The antecedents of leadership development practices in SMEs: The influence of HRM strategy and practice. *International Small Business Journal*, 34(6), 870-890.
- Gaskell, G., Allum, N., Wagner, W., Kronberger, N., Torgersen, H., Hampel, J. & Bardes, J. (2004). GM foods and the misperception of risk perception. *Risk Analysis*, 24(1), 185-194.
- Gimmler, A, Højme, P., Lautrup K. & Jakob, B. (2023). The Varieties of Applied Philosophy: Introduction. *Danish Yearbook of Philosophy*, 56(2), 105-111.
- Godin, B. (2012). Innovation and Conceptual Innovation in Ancient Greece. Chaire Fernand Dumont sur la Culture Project on the Intellectual History of Innovation. Working Paper No. 12 https://www.chairefernanddumont.ucs.inrs.ca/wp-content/uploads/2013/09/GodinB_2012_Innovation_and_Conceptual_Innovation.pdf (Letöltve 2024.03.11)
- Godin, B. (2014a). *The Politics of Innovation: Machiavelli and Political Innovation, or, How to Stabilize a Changing World*. Project on the Intellectual History of Innovation Working (17).
- Godin, B. (2014b). Invention, diffusion and linear models of innovation: the contribution of anthropology to a conceptual framework. *Journal of Innovation Economics & Management*, 15, 11-37.
- Godin, B. (2015a). *Innovation contested: the idea of innovation over the centuries*. Routledge, New York
- Godin, B. (2015b). Innovation: A Study in the Rehabilitation of a Concept. *Contributions to the History of Concepts*, 10(1), 45–68.

- Godin, B. (2015c). Innovation: A Study in the Rehabilitation of a Concept. *Contributions to the History of Concepts*, 10(1), 45–68.
- Godin, B. (2017). A conceptual history of innovation. In *The Elgar Companion to Innovation and Knowledge* (pp. 25-32). Edward Elgar Publishing.
- Godin, B. (2020a). *The Idea of Technological Innovation: A Brief Alternative History*. Cheltenham, Edward Elgar Publishing.
- Godin, B. (2020b). The English Reformation and the Invention of Innovation. Working Paper. INRS - Urbanisation Culture Société, Montréal. https://espace.inrs.ca/id/eprint/12550/1/Godin_2020.pdf (Letöltve 2024.03.11)
- Godin, B. (2021). Theories of innovation. In B. Godin, G. Gaglio & D. Vinck (Eds.), *Handbook on Alternative Theories of Innovation*. Edward Elgar, Cheltenham.
- Godin, B., Gérald, G. & Dominique, V. (2021). Introduction to the Handbook on Alternative Theories of Innovation. In B. Godin, G. Gaglio & D. Vinck (Eds.), *Handbook on Alternative Theories of Innovation*. Edward Elgar, Cheltenham.
- Good, I. J. (1965). Speculations concerning the first ultraintelligent machine. In F. L. Alt & M. Rubinoff (Eds.), *Advances in computers* (Vol. 6, pp. 31-88). Academic Press.
- Griffel, F. (2009). *Al-Ghazālī's philosophical theology*. Oxford University Press.
- Guido, F. (1988) Technical and technological innovation in Marx. *History and Technology*, 6(4), 299-324.
- Gulyás T., Szendrő P., Szécsi G. (2021). Innováció, etika, felelősség az információ korában. Adalékok az innovációetika fogalmi megalapozásához. *Kultúratudományi Szemle* 2021/3., 5-24.
- Guston, D. H. (2014). Understanding 'anticipatory governance'. *Social Studies of Science*, 44(2), 218-242.
- Guzzo, A. (1956). Les Études philosophiques. *Nouvelle Série, Psychanalyse*, 11(4), 665-667.
- Habuka, H., & Socol de la Osa, D. U. (2024). Shaping Global AI Governance: Enhancements and Next Steps for the G7 Hiroshima AI Process. Center for Strategic and International Studies.
- Heidegger, M. (1977). The Question Concerning Technology. In *The Question Concerning Technology and Other Essays*. Harper & Row.
- Heidegger, M. (1994). Kérdés a technika nyomán. In J. A. Tillmann (Ed.), *A későújkor józansága II. Olvasókönyv a tudományos-technikai világfelszámolás köréből* (pp. 117-133). Göncöl kiadó.
- Holman, C. (2018). *Machiavelli and the Politics of Democratic Innovation*. University of Toronto Press.
- Howard, J. A. & Sheth, J. N. (1969). *The theory of buyer behaviour*. Wiley, New York.
- Jasanoff, S. (2016). *The ethics of invention: Technology and the human future*. W. W. Norton & Company.
- Jerome, J. W. (2017). Big data: Catalyst for a privacy conversation. *Indiana Law Review*, 50(2), 643-671.
- Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
- Jonas, H. (1976). Responsibility Today: The Ethics of an Endangered Future. *Social Research*, 43(1), 77–97.

- Jonas, H. (1984). *The Imperative of Responsibility. In Search of an Ethics for the Technological Age*. Chicago & London, University of Chicago Press.
- Jonas, H. (2005). Az emberi cselekvés megváltozott természete. In L. András & B. Jávorski (Eds.), *Környezet és etika* (pp. 25–37). L'Harmattan.
- Karikó, S. (2002). Mi az alkalmazott filozófia? In S. Karikó (Ed.), *Az alkalmazott filozófia esélyei*. Áron kiadó.
- Kaufman, A. S. (2006). Responsibility, moral and legal. In D. M. Borchert (Ed.), *The Encyclopedia of Philosophy* (2nd edition) (pp. 444-451). Thomson Gale.
- Király, V. I. (2002). Alkalmazott filozófia és a filozófia szakmája. In K. Veress (Ed.), *A filozófia alkalmazása-alkalmazott filozófia*. A Babes-Bolyai Tudományegyetem Filozófia Tanszékcsoportja magyar tagozatán tartott tanévnyitó konferencia előadásai. 2001. november 3. Kolozsvár: Pro Philosophia.
- Király, V. I. (2004). Alkalmazott filozófia és a filozófia szakmája. Időszerűtlen aktualitások In V. I. Király (Ed.), *Kérdő Jelezés – több-csendbeni alkalmazott filozófiai zajhíborítás a szabad(ság) kérdés(é)ben*. Kalligram.
- Kolias, C., Kambourakis, G., Stavrou, A. & Voas, J. (2017). DDoS in the IoT: Mirai and other botnets. *Computer*, 50(7), 80-84.
- Kruk, R. (2011). The Vital Roots of European Enlightenment: Ibn Tufayl's Influence on Modern Western Thought, by Samar Attar. *Middle Eastern Literatures*, 14(1).
- Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Viking.
- LaFollette, H. (2003, Ed.). *The Oxford handbook of practical ethics*. Oxford University Press.
- Lányi, A. (2007). *A fenntartható társadalom*. L'Harmattan.
- Lányi, A. (2012). Az emberismeret és etika tantárgy elméleti hátteréről. *Új Pedagógiai Szemle*, 2012(4-6), 56-77.
- Leget, C., Borry, P. & de Vries, R. (2009). 'Nobody tosses a dwarf!' The relation between the empirical and the normative reexamined. *Bioethics*, 23(4), 226–235.
- Lennerfors, T. T., & Murata, K. (2023). Innovation ethics. In A. Rehn & A. Örtengren (Eds.), *Debating innovation* (pp. 189-210). Palgrave.
- Lévinas, E. (1999). *Teljesség és végtelen*. Tarnay L.(fordító). Jelenkor Kiadó.
- Lévinas, E. (2006). *Otherwise Than Being, Or, Beyond Essence*. Duquesne University Press.
- Lévy-Bruhl, L. (1884). *L'idée de responsabilité*. Hachette.
- Lippert-Rasmussen, K. (2016). The Nature of Applied Philosophy. In K. Lippert-Rasmussen, K. Brownlee & D. Coady (Eds.), *A Companion to Applied Philosophy* (pp. 1-17). Malden, Wiley.
- Lundvall, B-A. (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. Printer, London.
- Machiavelli, N. (1979). Niccolò Machiavelli művei, 1–2. Iványi, N. et al. (fordító). Európa, Budapest
- Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P. & Dewhurst, M. (2017). *Jobs lost, jobs gained: Workforce transitions in a time of automation*. McKinsey Global Institute.
- Maróth, M. (2007). Erkölcsfilozófia az iszlámban. *Iskolakultúra*, 17(3), 134–141.
- Maróth, M. (2013). Erőszak és térítés Alfarabi politikai filozófiájában. In B. Déry (Ed.), *Conversio* (157–162). ELTE BTK.

- Martin, D. (2007). Responsibility: A Philosophical Perspective. In G. Dewsbury & J. Dbozon (Eds.), *Responsibility and Dependable Systems* (pp. 21-42). Springer, London.
- Martins, H. (2022). The God of Artefacts: Vico's Principle and Technology. In H. M. Jerónimo (Ed.), *Portuguese Philosophy of Technology: Legacies and contemporary work from the Portuguese-Speaking Community* (pp. 25-52). Springer Verlag.
- Martins, H. (2023). The God of Artefacts: Vico's Principle and Technology. In H. M. Jerónimo (Ed.), *Portuguese Philosophy of Technology. Philosophy of Engineering and Technology*. Springer, Cham.
- Marx, K. (1955). *A tőke. A politikai gazdaságtan bírálata*. I. köt. 1. könyv. Rudas, L. & Nagy, T. (fordító). Szikra, Budapest
- Marx, K. (1972). A gépekről szóló töredék. In *A politikai gazdaságtan bírálata* alapvonalai. II. kötet. Kossuth kiadó.
- Mattila, J. (2017). The Ethical Progression of the Philosopher in Al-Rāzī and Al-Fārābī. *Arabic Sciences and Philosophy*, 27(1), 115–137.
- Maurer, B., Nelms, T. C. & Swartz, L. (2013). When perhaps the real problem is money itself! The practical materiality of Bitcoin. *Social Semiotics*, 23(2), 261-277.
- McAfee, A. & Brynjolfsson, E. (2017). *Machine, platform, crowd: Harnessing our digital future*. W. W. Norton & Company.
- McKeon, R. (1957). The Development and the Significance of the Concept of Responsibility. *Revue Internationale de Philosophie*, 11(39) 3-32. Entretiens De Paris (13-15 september 1956) Institut International de Philosophie.
- McKeon, Z. K. (1998). General Introduction. In K. McKeon, R. McKeon, Z. Karl & W. G. Swenson (Eds.), *Selected writings of Richard Mckeon* (pp. 1-21). University of Chicago Press.
- Meier, C. (1990). *The Greek Discovery of Politics*. Cambridge, Mass.
- Merton, R. K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.
- Miller, F. G. & Wertheimer, A. (2010). Preface to a theory of consent transactions: Beyond valid consent. In *The ethics of consent: Theory and practice* (pp. 79-105). Oxford University Press.
- Mitcham, C. (1994). *Thinking through Technology: The Path between Engineering and Philosophy*. The University of Chicago Press.
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501-507.
- Morris, N. (2024). Ethics of innovation: A framework for responsible innovation governance. *SMU Science & Technology Law Review*, 27, 39.
- Morton, J. A. (1968). The Innovation of Innovation. *IEEE Transactions on Engineering Management EM*, 15(2), 57-65.
- Mulgary, G. & Albury, D. (2003). *Innovation in the Public Sector*. Strategy Unit, Cabinet Office.
- Mumford, L. (1961). *The city in history: its origins, its transformations, and its prospects*. Harcourt Brace Jovanovich.
- Muresan, V. & Majima, S. (2013). Is Applied Ethics Still an Application of Ethics? A Short History of Applied Ethics in Romania. In V. Muresan & S. Majima (Eds.), *Applied*

- Ethics: Perspectives from Romania*. Center for Applied Ethics and Philosophy, Hokkaido University.
- Narayanan, A., Bonneau, J., Felten, E., Miller, A. & Goldfeder, S. (2016) Bitcoin and cryptocurrency technologies. Princeton University Press.
- Nasr, S. H. (1968). *Science and Civilization in Islam*. Mass. Harvard University Press.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Department of Health and Human Services.
- Nietzsche, F. (1996). *Adalék a morál genealógiájához*. Holnap Kiadó.
- Nietzsche, F. (2004). *Így szólott Zarathustra*. Osiris.
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32-48.
- Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the internet worldwide*. Cambridge University Press.
- Nyíri, K. (1996). The Humanities in the Age of Post-Literacy. *Budapest Review of Books - English Edition*, 6(3), 110-116.
- Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E. & Guston, D. (2013). A framework for responsible innovation. In R. Owen, J. Bessant & M. Heintz (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 27-50). John Wiley & Sons.
- Owens, W. A., Dam, K. W. & Lin, H. (2019). *Technology, policy, law, and ethics regarding U.S. acquisition and use of cyberattack capabilities*. National Academies Press.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Paterson, T. (1989). The Secular Control of Scientific Power in the Political Philosophy of Francis Bacon. *Polity*, 21(3), 457–480.
- Pittz, S. F. & Pittz, T. G. (2024). Towards an ethical awareness of entrepreneurs: a Nietzschean perspective on creative destruction. *Entrepreneurship & Regional Development*, 36(5–6), 707–721.
- Platón (1984). *Platón összes művei I–III*. Európa Kiadó.
- Pocock, J. G. A. (1975). *The Machiavellian Moment*. Princeton University Press.
- Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum*, (2), 79.
- Reijers, W., Wright, D., Brey, P., Weber, K., Rodrigues, R., O’Sullivan, D. & Gordijn, B. (2017). Methods for practising ethics in research and innovation: A literature review, critical analysis and recommendations. *Science and Engineering Ethics*.
- Reinert, E. S. (2019). Creative Destruction in Economics: Nietzsche, Sombart, Schumpeter (with Hugo Reinert). In R. Kattel (Ed.), *The Visionary Realism of German Economics: From the Thirty Years’ War to the Cold War* (pp. 385–412). Anthem Press.
- Reinert, H. & Reinert, E. S. (2006). Creative Destruction in Economics: Nietzsche, Sombart, Schumpeter. In J. G. Backhaus & W. Drechsler (Eds.), *Friedrich Nietzsche (1844–1900). The European Heritage in Economics and the Social Sciences*, vol 3. Springer.
- Richards, N. M. (2013). The dangers of surveillance. *Harvard Law Review*, 126(7), 1934-1965.
- Roco, M. C. & Bainbridge, W. S. (2003). Overview converging technologies for improving human performance. In M. C. Roco & W. S. Bainbridge (Eds.), *Converging technologies for improving human performance* (pp. 1-27). Springer.

- Roessler, B. (2005). *The value of privacy*. Polity Press.
- Rogers, E. M. (1983). *Diffusion of Innovation*. (3rd edition). Free Press.
- Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J. & Bengio, Y. (2019). Tackling climate change with machine learning. *arXiv preprint arXiv:1906.05433*.
<https://doi.org/10.48550/arXiv.1906.05433>
- Rorty, R. (2006). Is Philosophy Relevant to Applied Ethics? *Bus Ethics Q*, 16(3), 369–380.
- Routledge encyclopedia of philosophy (1998). Edward Craig (Ed.). 'responsibility' szócikk. Routledge.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215.
- Russell, S. J. & Norvig, P. (2009). *Artificial intelligence: A modern approach*. Prentice Hall.
- Sand, M. (2018). A vision for the future of responsibility. In M. Sand (Ed.), *Futures, visions, and responsibility* (pp. 15-30). Springer VS.
- Santarelli, E. & Pesciarelli, E. (1990). The Emergence of a Vision: The Development of Schumpeter's Theory of Entrepreneurship. *History of Economic Policy*, 22(4).
- Sarewitz, D. (2016). Saving science. *The New Atlantis*, 49, 4-40.
- Scharre, P. (2018). *Army of none: Autonomous weapons and the future of war*. W. W. Norton & Company.
- Schneier, B. (2015). *Data and Goliath: The hidden battles to collect your data and control your world*. W. W. Norton & Company.
- Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Interest and the Business Cycle*. Transaction Publishers.
- Schumpeter, J. A. (1939). *Business Cycles: A theoretical, Historical and Statistical Analysis of the Capitalist Process*. McGraw-Hill Book Company.
- Schumpeter, J. A. (1975). *Capitalism, Socialism and Democracy*. Harper. 82-85.
- Schumpeter, J. A. (1980). *A gazdasági fejlődés elmélete*. Közgazdasági és Jogi Kiadó.
- Schwab, K. (2016). *The Fourth Industrial Revolution*. World Economic Forum.
- Selinger, W. (2018). The Forgotten Philosopher: A Review Essay on Richard McKeon. *The Review of Politics*, (80), 137-150.
- Senn, P. R. (2006). The Influence of Nietzsche on the History of Economic Thought. In J. G. Backhaus & W. Drechsler (Eds.), *Friedrich Nietzsche (1844–1900). The European Heritage in Economics and the Social Sciences*, vol 3. Springer.
- Singer, P. (2011). *Practical Ethics* (3rd edition). Cambridge University Press.
- Smith, R. E. & Brooks, D. J. (2018). *Security science: The theory and practice of security*. Butterworth-Heinemann.
- Society for Applied Philosophy (2024). SAP. <https://www.appliedphil.org/>
- Stankovic, J. A. (2014.) Research directions for the Internet of Things. *IEEE Internet of Things Journal*, 1(1), 3-9.
- Stephan, P. (2012). *How economics shapes science*. Harvard University Press.
- Stevenson, L. (1970). Applied Philosophy. *Metaphilosophy*, 1(3), 258–267.

- Stilgoe, J., Owen, R. & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568-1580.
- Studer, H. D. (1998). Francis Bacon on the Political Dangers of Scientific Progress. *Canadian Journal of Political Science*, 31(2), 219–234.
- Sugarman, J. (2004). The future of empirical research in bioethics. *Journal of Law, Medicine & Ethics*, 32(2), 226–231.
- Szécsi, G. (2002). A filozófiai nyelv alkalmazhatósága. In S. Karikó (Ed.), *Az alkalmazott filozófia esélyei*. Áron kiadó.
- Szilágyi, T. (2021). A felelősség társadalma. *Kultúratudományi Szemle*, 3(4), 38-45.
- Szilágyi, T. (2022). A felelősség kérdése a kortárs etikai diskurzusban. In G. Szécsi, J. Boros & Zs. Nemeskéri (Eds.), *XXI. századi narratívák: Tanulmányok a filozófia, művészet és gazdaság metszéspontján* (pp. 111-122). Belvedere Meridionale.
- Szilágyi, T. (2023). A felelősség fogalma Richard Mckeon filozófiai gondolkodásában. In E. Daróczi & S. Laczkó (Eds.), *A felelősség* (pp. 385-389). Státus Könyvkiadó, Pro Philosophia Szegedi Alapítvány, Magyar Filozófiai Társaság.
- Szilágyi, T. (2024a). Innováció vs. fenntarthatóság: A kortárs tudományos gyakorlat etikai dilemmái. In F. Dudok (Ed.), *Környezeti nevelés a 21. században: Tapasztalatok és jó gyakorlatok* (pp. 142-150). Pécsi Tudományegyetem.
- Szilágyi, T. (2024b). A mesterséges intelligencia alkalmazásának etikai vonatkozásai az oktatási innovációban. In T. K. Fodoroné (Ed.), *Felsőoktatási Lifelong Learning a képességfejlesztés, a kultúrákövetítés és fenntarthatóság tükrében. A hatékony és innovatív tudástranszferek lokális/regionális és globális trendjei* (pp. 209-214). MELearn Egyesület.
- Tapscott, D. & Tapscott, A. (2016) *Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world*. Penguin.
- Taylor, A. E. (1997). *Platón*. Osiris könyvtár.
- Tenner, E. (1996). *Why things bite back: Technology and the revenge of unintended consequences*. Knopf.
- Thacker, E. (2011). *In the Dust of This Planet: Horror of Philosophy, 1*. Zero Books.
- Tóth, I. J. (2002). Az alkalmazott filozófia státuszáról. In S. Karikó (Ed.), *Az alkalmazott filozófia esélyei*. Áron kiadó.
- Tóth, I. J. (2007). Hans Jonas életéről és munkásságáról. *Különbség*, 9(1), 51-57.
- Tufajl, A. B. I. (1961). *A természetes ember*. Európa Könyvkiadó.
- van de Poel, I. & Royakkers, L. (2007). The ethical cycle. *Journal of Business Ethics*, 71(1), 1-13.
- Van de Poel, I. & Royakkers, L. (2011). *Ethics, technology, and engineering: An introduction*. Wiley-Blackwell.
- van Eck, N. J. & Waltman, L. (2014). Visualizing bibliometric networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring scholarly impact* (pp. 285-320). Springer.
- Vanyó, L. (1970). Az újkori gondolkodás és Aquinói Szent Tamás. *Vigilia*, (35), 155–159.
- Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.
- Verbeek, P. P. (2015). Beyond interaction: A short introduction to mediation theory. *Interactions*, 22(3), 26-31.

- Verbeek, P. P. (2016). Toward a theory of technological mediation: A program for postphenomenological research. In J. K. Berg O. Friis & R. C. Crease (Eds.), *Technoscience and postphenomenology: The Manhattan papers* (pp. 189-204). Lexington Books.
- Verbeek, P. P. (2017). Technology mediation: How technology organizes human-world relations. *Fatum*, 6(S), 39.
- Vico, G. (1710). *De antiquissima italorum sapientia, ex Linguae Latinae Originibus eruenda*. Libri tres Joh. Baptistae a Vico Neapolitani Regii Eloquentiae Professoris. Mosca.
- Vico, G. (1953). *Az új tudomány*. Dienes, G. (fordító). Akadémiai Kiadó.
- von Schomberg, R. (2011). Prospects for technology assessment in a framework of responsible research and innovation. In M. Dusseldorp & R. Beecroft (Eds.), *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden*. Wiesbaden, Germany: Vs Verlag.
- von Schomberg, R. (2013). A vision of responsible research and innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible innovation: Managing the responsible emergence of science and innovation in society* (pp. 51-74). John Wiley & Sons.
- Voss, P. (2007). Essentials of general intelligence: The direct path to artificial general intelligence. In B. Goertzel & C. Pennachin (Eds.), *Artificial general intelligence* (pp. 131-157). Springer.
- Walsh, S. (2023). Marx, subsumption and the critique of innovation. *Organization*, 30(2), 345-360.
- Wang, R. H. & Bickenbach, J. (2021). Ethics of health research and innovation. In A. Sixsmith, J. Sixsmith, A. Mihailidis, & M. L. Fang (Eds.), *Knowledge, innovation, and impact: A guide for the engaged health researcher* (pp. 455-474). Springer.
- Wanzenböck, I. Wesseling, J. H., Frenken, K. Hekkert, M. & Weber, M. (2020). A framework for mission-oriented innovation policy: Alternative pathways through the problem–solution space, *Science and Public Policy*, 47(4), 474–489.
- Wardle, C. & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe report.
- Warren, B. (1992). Back to Basics: Problems and Prospects for Applied Philosophy. *Journal of Applied Philosophy*, 9(1), 13–19.
- Watson, G. (1996). Two Faces of Responsibility. *Philosophical Topics*, 24(2), 227–248.
- Weber, M. (2014). *A tudomány és a politika mint hivatás*. Kossuth Kiadó.
- Weisz, G. (1990). Introduction. In G. Weisz (Ed.), *Social science perspectives on medical ethics* (pp. 3–15). Kluwer.
- Wendling, A. E. (2009). *Karl Marx on technology and alienation*. Palgrave-Macmillan.
- Yeung, K. (2017). 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118-136.
- Yudkowsky, E. (2013). Intelligence explosion microeconomics. *Machine Intelligence Research Institute, Technical Report 2013-1*.
- Zohar, A. (2015). Bitcoin: Under the hood. *Communications of the ACM*, 58(9), 104–113.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Internetes forrás

- Az Európai Parlament állásfoglalása (2020. október 20.) a Bizottsághoz intézett ajánlásokkal a mesterséges intelligencia, a robotika és a kapcsolódó technológiák etikai szempontjainak keretéről (2020/2012(INL)). HL C 404, 2021.10.6., 4-20. <https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52020IP0275>. Letöltés: 2024. június 3.
- European Commission: Directorate-General for Research and Innovation, & Schomberg, R. (2011). *Towards responsible research and innovation in the information and communication technologies and security technologies fields* (R. Schomberg, Ed.). Publications Office. <https://op.europa.eu/en/publication-detail/-/publication/60153e8a-0fe9-4911-a7f4-1b530967ef10> Letöltés: 2024. június 3.
- European Commission: Directorate-General for Research and Innovation. (2013). *Options for strengthening responsible research and innovation: Report of the expert group on the state of art in Europe on responsible research and innovation*. Publications Office. <https://op.europa.eu/en/publication-detail/-/publication/1e6ada76-a9f7-48f0-aa86-4fb9b16dd10c> Letöltés: 2024. június 3.
- Forgeard, V. (2023). The vital role of ethics in shaping the future of innovation. *Brilliantio*. <https://brilliantio.com/why-should-ethics-be-a-part-of-invention-and-discovery/> Letöltés: 2024. augusztus 3.
- Future of Life Institute. (2017). Asilomar AI principles. <https://futureoflife.org/ai-principles/>. Letöltés: 2024. március 5
- G7 Hiroshima AI Process. (2023). Hiroshima Process International Guiding Principles for All AI Actors. <https://www.mofa.go.jp/files/100573471.pdf>. Letöltés: 2024. augusztus 3.
- G7 Hiroshima AI Process. (2023). Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems. <https://www.mofa.go.jp/files/100573473.pdf>. Letöltés: 2024. augusztus 3.
- High-Level Expert Group on Artificial Intelligence (2019) Ethics guidelines for trustworthy AI. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Letöltés: 2024. március 5
- Leslie, D. (2021) Understanding Artificial Intelligence Ethics and Safety: A Guide for the Responsible Design and Implementation of AI Systems in the Public Sector (June 10, 2019). https://www.turing.ac.uk/sites/default/files/2019-08/understanding_artificial_intelligence_ethics_and_safety.pdf. Letöltés: 2024. március 13.
- Lorenz, P., K. Perset and J. Berryhill (2023), Initial policy considerations for generative artificial intelligence, OECD Artificial Intelligence Papers, No. 1, OECD Publishing, Paris, https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en. Letöltés: 2024. augusztus 3.
- National Institutes of Health Office of Intramural Research. (2024). *Annual review of ethics case studies*. <https://oir.nih.gov/sourcebook/ethical-conduct/responsible-conduct-research-training/annual-review-ethics-case-studies> Letöltés: június 3.
- OECD (2024) Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Letöltés: 2024. augusztus 3.

United Nations Educational, Scientific and Cultural Organization. (2021). Recommendation on the ethics of artificial intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000380455> Letöltés: 2024. március 5

World Economic Forum. (2020). Reskilling revolution: Better skills, better jobs, better education for a billion people by 2030. World Economic Forum. <https://initiatives.weforum.org/reskilling-revolution/home> Letöltés: 2024. január 5.

Ábrák jegyzéke

1. ábra. Az alkalmazott etika hétosztatú modellje (saját szerkesztés).....	21
2. ábra. Az innovációetika és az innováció etikája kulcsszavak előfordulása (db) (forrás: Scopus).....	83
3. ábra. A megjelent tanulmányok tudományterületi megoszlása (innovációetika) (forrás: Scopus).....	84
4. ábra. A publikációk megoszlása évek és kategória szerinti bontásban (db) 2013-2024 (saját szerkesztés) (forrás: Scopus).....	85
5. ábra. Publikációk megoszlása földrajzi megoszlás szerint – az első 5. ország (db) (etika/erkölcs és felelősség) (forrás: Scopus)	86
6. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és felelősség) (forrás: Scopus).....	87
7. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és innováció) (forrás: Scopus).....	88
8. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és felelősség és innováció) (forrás: Scopus).....	88
9. ábra. Az innováció és az AI kulcsszavak előfordulása (db) (forrás: Scopus).....	90
10. ábra. A megjelent tanulmányok tudományterületi megoszlása (etika/erkölcs és innováció) (forrás: Scopus).....	90
11. ábra. A publikációk megoszlása évek és kategória szerinti bontásban (db) 2013-2024 (saját szerkesztés) (forrás: Scopus).....	91
12. ábra. A felelős döntéshozatali mátrix folyamatábrája (saját szerkesztés)	148
13. ábra. A példa responzibilitási indexének radardiagramja (saját szerkesztés)	154

Táblázatok jegyzéke

1. táblázat. A felelős kutatás és innováció jellemzőinek áttekintése (Von Schomberg, 2011) .	79
2. táblázat. A publikációk megoszlása évek és kategória szerinti bontásban (db) 2013-2024 (saját szerkesztés) (forrás: Scopus).....	85
3. táblázat. A SCOPUS adatbázisban szereplő dokumentumok nyelvi megoszlása (saját szerkesztés).....	87
4. táblázat. A megjelent tanulmányok tudományterületi megoszlása az első hat kategóriában (forrás: Scopus).....	89
5. táblázat. A megjelent tanulmányok tudományterületi megoszlása (forrás: Scopus)	92
6. táblázat. Legfontosabb alapelvek és fókuszterületek (saját szerkesztés).....	110
7. táblázat. Innováció és felelősség a dokumentumokban (saját szerkesztés).....	112
8. táblázat. Felelős döntéshozatali mátrix kérdései (saját szerkesztés)	142
9. táblázat. Példa az A etikai dimenzió kiszámításához (saját szerkesztés)	145
10. táblázat. Példa a responzibilitási index (Ri) kiszámításához (saját szerkesztés)	145
11. táblázat. Példa Ri pontszámok az iRi érték kiszámításához (saját szerkesztés).....	146
12. táblázat. Az iRi értéktartományok határai (saját szerkesztés)	147
13. táblázat. Az emberi élet és méltóság tisztelete- pontozás (saját szerkesztés).....	151
14. táblázat. A magánélet és adatvédelem - pontozás (saját szerkesztés).....	151
15. táblázat. Diszkrimináció és előítéletmentesség (saját szerkesztés)	152
16. táblázat. Elszámoltathatóság és felelősség - pontozás (saját szerkesztés).....	152
17. táblázat. Társadalmi és környezeti hatás - pontozás (saját szerkesztés).....	153
18. táblázat. Kulturális integritás és identitás (saját szerkesztés)	153

Mellékletek

I. Melléklet

A SCOPUS adatbázis elemzésének kiegészítő adatai, az összes táblázat és grafikon saját szerkesztés.

I.M.-19. táblázat. A SCOPUS adatbázisban használt keresési karakterlánc

	Keresési karakterlánc (string)	Keresési időszak	Találati eredmény
Etika/erkölcs és felelősség	TITLE-ABS-KEY(("ethics"OR"moral")AND("responsib*"))AND PUBYEAR>1995ANDPUBYEAR<2025AND(LIMIT-TO(EXACTKEYWORD,"Ethics"))AND(LIMIT-TO(DOCTYPE,"ar")ORLIMIT-TO(DOCTYPE,"cp")ORLIMIT-TO(DOCTYPE,"ch")ORLIMIT-TO(DOCTYPE,"bk"))	1996-2024	10567 db
Etika/erkölcs és innováció	TITLE-ABS-KEY(("ethics"OR"moral")AND("innovation"))ANDPUBYEAR>1995ANDPUBYEAR<2025AND(LIMIT-TO(EXACTKEYWORD,"Ethics")ORLIMIT-TO(EXACTKEYWORD,"Innovation"))AND(LIMIT-TO(DOCTYPE,"ar")ORLIMIT-TO(DOCTYPE,"cp")ORLIMIT-TO(DOCTYPE,"ch")ORLIMIT-TO(DOCTYPE,"bk"))	1996-2024	2369 db
Etika/erkölcs és felelősség és innováció	TITLE-ABS-KEY(("ethics"OR"moral")AND("responsib*")AND("innovation"))ANDPUBYEAR>1995ANDPUBYEAR<2025AND(LIMIT-TO(EXACTKEYWORD,"Ethics"))AND(LIMIT-TO(DOCTYPE,"ar")ORLIMIT-TO(DOCTYPE,"cp")ORLIMIT-TO(DOCTYPE,"ch")ORLIMIT-TO(DOCTYPE,"bk"))	1996-2024	458 db
Innováció etika/az innováció etikája	TITLE-ABS-KEY("innovation ethics"OR"ethics of innovation")AND PUBYEAR>1995 AND PUBYEAR<2025	1996-2024	72 db.

I.M.-2. táblázat. A publikációk földrajzi megoszlása

Ország/Terület	Dokumentumok száma (db)		
	Etika/erkölcs és felelősség	Etika/erkölcs és innováció	Etika/erkölcs és felelősség és innováció*
Egyesült Államok	3347	643	111
Egyesült Királyság	1425	373	96
Kanada	677	163	23**
Ausztrália	676	141	31
Hollandia	469	160	52

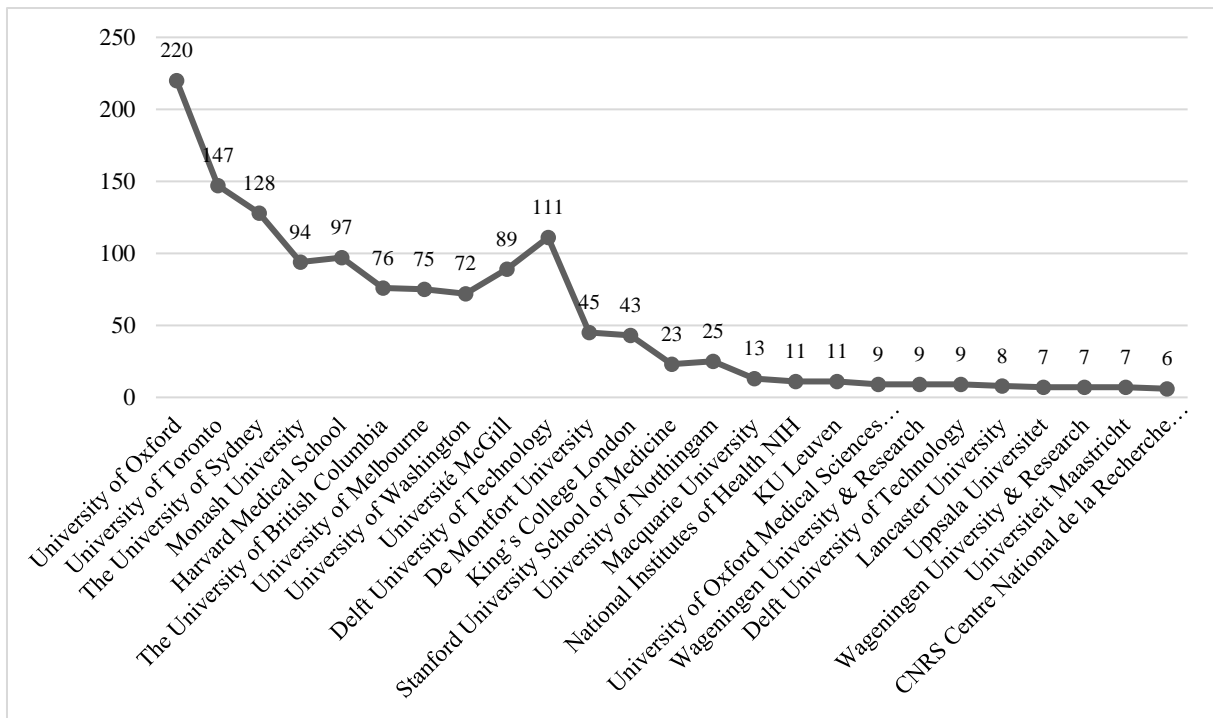
* az adott kategóriában az első 5 hely között szerepel még Németország 30 db dokumentummal

** Kanada csak a 7. helyen áll, ezt megelőzi még Olaszország 28 db dokumentummal

I.M.-3. táblázat. A talált publikációk kategória és egyetemi szintű bontása (db)

Egyetemi affiliáció	Dokumentumok száma (db)			
	Etika/erkölcs és felelősség	Etika/erkölcs és innováció	Etika/erkölcs és felelősség és filozófia	Etika/erkölcs és felelősség és innováció
University of Oxford	151	33	22	14
University of Toronto	120	27		
The University of Sydney	94	23	11	
Monash University	82		12	
Harvard Medical School	79	18		
The University of British Columbia	76			
University of Melbourne	75			
University of Washington	72			
Université McGill	70	19		
Delft University of Technology	69	27		15
De Montfort University		25		20
King's College London		23	12	8
Stanford University School of Medicine		23		
University of Nottingham		19		6
Macquarie University			13	
National Institutes of Health NIH			11	
KU Leuven			11	
University of Oxford Medical Sciences Division			9	
Wageningen University & Research			9	
Delft University of Technology			9	
Lancaster University				8
Uppsala Universitet				7
Wageningen University & Research				7
Universiteit Maastricht				7
CNRS Centre National de la Recherche Scientifique				6

A publikációk számát egyetemenként összesítettem, így az első 10(-10) egyetemet és az általuk megjelentett publikációk száma látható a táblázatban.

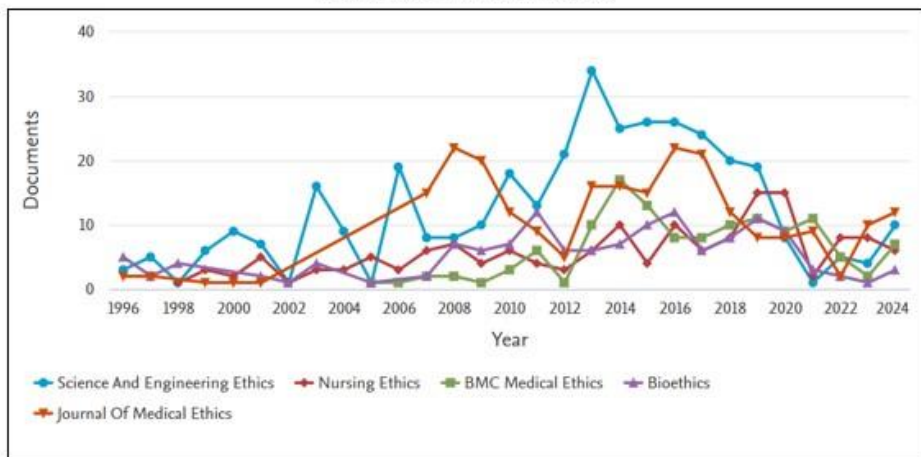


I.M.-1. ábra. A publikációk száma összesítve egyetemek szerinti bontásban (db)

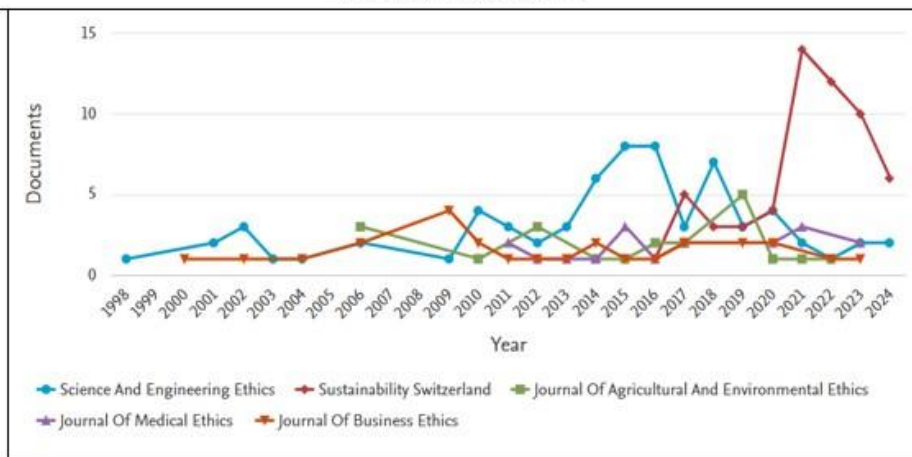
I.M.-4. táblázat. A publikációk tudományterületi szakfolyóiratok szerinti bontása (kategóriák, megjelent publikációk száma (db) 1996-)

Folyóirat neve	Dokumentumok száma (db)			
	Etika/erkölcs és felelősség	Etika/erkölcs és innováció	Etika/erkölcs és felelősség és filozófia	Etika/erkölcs és felelősség és innováció
Science And Engineering Ethics	357	69		38
Journal Of Medical Ethics	241	18	134	4
Nursing Ethics	154			
Bioethics	137	17	81	
BMC Medical Ethics	128			6
American Journal Of Bioethics	125			
Journal Of Business Ethics	112	26	36	
Sustainability Switzerland	108	58		16
Medicine Health Care And Philosophy	76		16	
Journal Of Agricultural And Environmental Ethics		21	34	9
Journal Of Law Medicine And Ethics		16		
Social Science And Medicine		16	28	
Technological Forecasting And Social Change		15		
Kennedy Institute Of Ethics Journal			34	
Journal Of Medicine And Philosophy			20	
AI And Society			16	
Conference On Human Factors In Computing Systems Proceedings				7
Journal Of Information Communication And Ethics In Society				6
Life Sciences Society And Policy				6
Journal Of Responsible Innovation				4

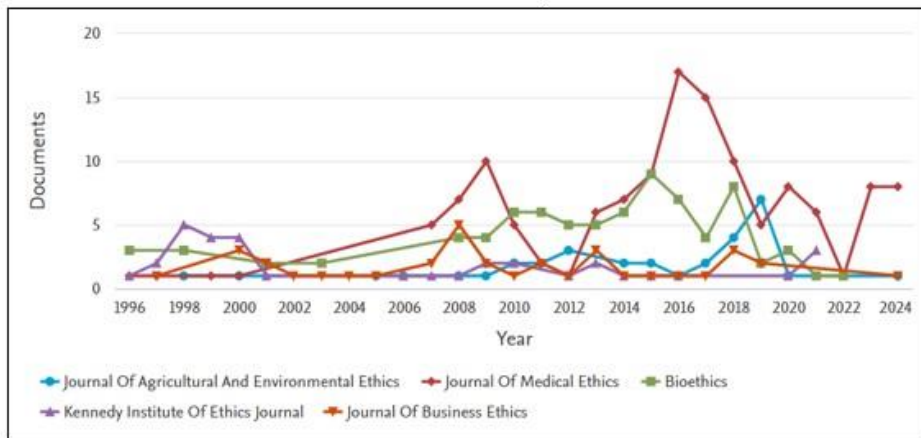
Etika/erkölcs és felelősség



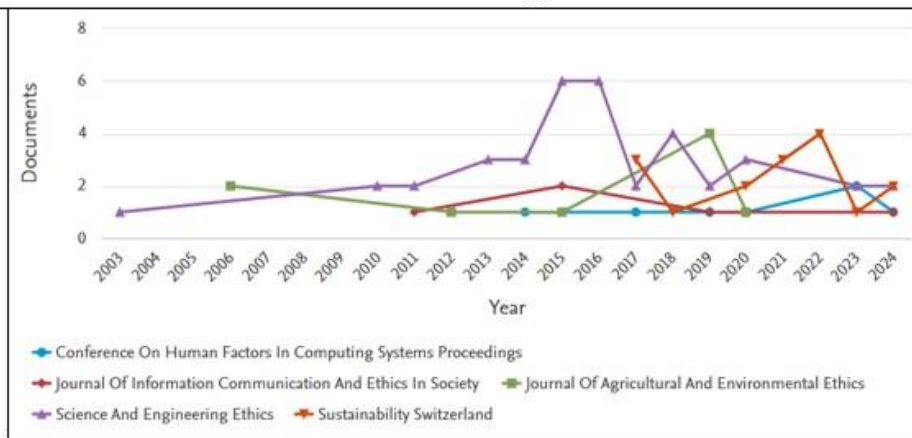
Etika/erkölcs innováció



Etika/erkölcs és felelősség és filozófia



Etika/erkölcs és felelősség és innováció



I.M.-2. ábra. Az öt vezető szakfolyóiratban megjelent publikációk számának megoszlása

A legfontosabb szerzők szakterületi megoszlásában

I.M.-5. táblázat. Etika-felelősség és etika-innováció területen legtöbbet publikáló szerzők neve, a publikációk száma és szakterülete

Etika/erkölcs és felelősség		
Szerző neve	Dokumentumok száma (db)	Szakterület megnevezése
McCullough, L. B.	46	orvosi etika
Chervenak, F. A.	34	orvosi etika
Bredenoord, A. L.	22	bioetika, orvosi etika
Resnik, D. B.	21	bioetika, orvosi etika, tudományfilozófia
Stahl, B. C.	20	a fejlődő technológia etikája, számítógépes etika
Pennings, G.	18	bioetika, orvosi etika,
Gastmans, C. A. E.	16	bioetika, orvosi etika,
Knoppers, B. M.	16	bioetika
Savulescu, J.	16	bioetika, orvosi etika
Sugarman, J.	14	bioetika, orvosi etika
Etika/erkölcs és innováció		
Szerző neve	Dokumentumok száma (db)	Szakterület megnevezése
Stahl, B. C.	17	mesterséges intelligencia és számítógépes etika, innovációs etika
Angelos, P.	8	orvosi etika
Bredenoord, A. L.	8	bioetika
Fins, J. J.	8	bioetika
Floridi, L.	7	mesterséges intelligencia, filozófiai szempontok, gépi tanulás
Lehoux, P. N.	7	bioetika, orvosi etika, mesterséges intelligencia
Lipworth, W. L.	7	bioetika
Ladikas, M.	6	technológiai etika
Mulej, M.	6	társadalometika
Osseweijer, P.	6	ökoetika

I.M.-6. táblázat. Etika-felelősség-filozófia és etika-felelősség-innováció területén legtöbbet publikáló szerzők neve, a publikációk száma és szakterülete

Etika/erkölcs és felelősség és filozófia		
Szerző neve	Dokumentumok száma (db)	Szakterület megnevezése
McCullough, L. B.	4	orvosi etika
Munthe, C.	4	bioetika
Pratt, B. F.	4	egészségkutatás, bioetika
Snyder, J. C.	4	bioetika, orvosi etika
Blok, V.	3	környezetetika
Buchanan, A. E.	3	politikai és jogi etika, bioetika
Crooks, V. A.	3	egészségügyi szolgáltatások etikája
Dabrock, P.	3	bioetika, az emberi életformák technikai és tudományos viszonyrendszerének etikája
Doris, J. M.	3	szervezetetika
Douglas, T. M.	3	alkalmazott etika, bioetika
Etika/erkölcs és felelősség és innováció		
Szerző neve	Dokumentumok száma (db)	Szakterület megnevezése
Stahl, B. C.	17	mesterséges intelligencia és számítógépes etika, innovációs etika
Flipse, S. M.	5	technológiaetika, innovációetika
Osseweijer, P.	6	ökoetika
Ulnicane, I.	5	technológiaetika, innovációetika
Fothergill, B. T.	4	bioetika, innovációetika
Jirotko, M.	4	technológiaetika, innovációetika
Knight, W.	4	technológiaetika, innovációetika
Pennings, G.	4	bioetika, orvosi etika
van der Sanden, M. C. A.	4	technológiaetika
Akintoye, S.	3	társadalometika, technológiaetika