

**THE PAST, PRESENT AND FUTURE OF
ARTIFICIAL INTELLIGENCE FROM THE
PERSPECTIVE OF SENIOR AND JUNIOR
EXPERTS (PART 2)****A MESTERSÉGES INTELLIGENCIA
MÚLTJA, JELENE ÉS JÖVŐJE A SENIOR ÉS
A JUNIOR SZAKÉRTŐK SZEMSZÖGÉBŐL
(2. RÉSZ)^{1,2}**HEITLERNÉ LEHOCZKY Mária³ – KOLLÁR Csaba⁴**Abstract**

Our research focused on the current state and possible future of artificial intelligence and its impact on society. In the first part, we described the methodology of the focus group research, which was identical in both cases, and the results of the senior expert group, and in the second part, after a detailed analysis of the junior expert focus group research, we compared the opinions of the two groups by means of comparative analysis. Finally, based on our results, we formulate practical recommendations for realising a human-centred future based on ethical AI.

Keywords

artificial intelligence, expert survey, research methodology, online focus group, Artificial Intelligence Workshop

Absztrakt

Kutatásunk a mesterséges intelligencia jelenlegi helyzetének és lehetséges jövőképeinek áttekintésére, társadalmi hatásaira irányult. Az első részben ismertettük a fókuszcsoportos kutatás módszertanát, mely mindkét esetben azonos volt, és a senior szakértői csoport eredményeit, ennek folytatásaként a második részben a junior szakértői fókuszcsoportos kutatás részletes elemzése után a két csoport véleményét vettük össze komparatív elemzéssel. Végül eredményeink alapján gyakorlati ajánlásokat fogalmaztunk meg a humánfókuszú, etikus mesterséges intelligenciára épülő jövő megvalósításához.

Kulcsszavak

mesterséges intelligencia, szakértői megkérdezés, kutatómódszertan, online fókuszcsoport, Mesterséges Intelligencia Műhely

¹ A tanulmány kutatási háttérének alapját a 2021-1-HU01-KA220-HED-000029536 azonosítószámú „HEDY – Life in the AI Era” című Erasmus+ pályázatban a nevezett szerzők által végzett fókuszcsoportos szakértői megkérdezés jelentette.

² A tanulmány első része a Biztonságtudományi Szemle 2022. évi, IV. évf. 1. számában jelent meg:

<https://biztonsagtudomanyi.szemle.uni-obuda.hu/index.php/home/article/view/208/184>

³ maria.lehoczky@gmail.com | ORCID: 0000-0003-0588-715X | PhD student, Óbuda University Doctoral School for Safety and Security Sciences | member, Óbuda University Bánki Donát Faculty of Mechanical and Safety Engineering Artificial Intelligence Workshop | doktorandusz, Óbudai Egyetem Biztonságtudományi Doktori Iskola | tag, Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar Mesterséges Intelligencia Műhely

⁴ kollar.csaba@uni-obuda.hu | ORCID: 0000-0002-0981-2385 | senior research fellow and leader, Óbuda University Bánki Donát Faculty of Mechanical and Safety Engineering Artificial Intelligence Workshop | tudományos főmunkatárs és vezető, Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar Mesterséges Intelligencia Műhely

A JUNIOR SZAKÉRTŐI CSOPORT

A csoport bemutatása

Az alapsokaságot az Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar gépészmérnöki, mechatronikai mérnöki és biztonságtechnikai mérnök szakos hallgatói alkották. Az alapsokaságon belül egy kisebb csoportot képeztünk, melybe azok tartoznak, akik felvették 2022. tavaszi a félévében a Kollár Csaba által meghirdetett „Mesterséges intelligencia a műszaki életben” című tárgyat. A tantárgyat felvett hallgatók számára küldtük ki email-ben a felkérést, hogy lenne-e kedvük részt venni egy, a mesterséges intelligencia kihívásaival, lehetőségeivel, kockázataival, s jövőjével kapcsolatos fókuszcsoporthoz beszélgetésen. Felkérésünkre 5 nappali tagozatos hallgató jelentkezett: Bárczi Dávid, Barna Bianka Rita, Bayaraa Burtejin, Benkő Noémi, Fábíán Zsolt. A junior szakértői fókuszcsoporthoz így 3 nő és 2 férfi vett részt, a legfiatalabb résztvevő 20, a legidősebb 28 éves volt, az átlagéletkoruk 24 év. Mind az öt junior szakértőnk jelenleg biztonságtechnikai mérnök alapszakos hallgató, információbiztonsági specializáción/szakirányon. Előtanulmányaik között a következő képzéseket nevesítették: logisztikai ügyintéző, tűzvédelmi előadó, egészségügyi szakközépiskola, gimnázium, hálózati rendszergazda, gépész, gépjárműtechnológus, CNC programozó, rendszerüzemeltető.

A csoport körében vizsgált kérdések, területek

Nyitókör:

- Kérek egy rövid bemutatkozást: keresztnév, életkor, milyen karra/szakra jársz?

Ráhangelés:

Gyerekkorotokból milyen emlékeket tudtok felidézni a mesterséges intelligenciával, illetve a robotokkal kapcsolatban?

- Volt valamilyen robotos játékotok, vagy láttatok valamilyen robotokról, mesterséges intelligenciáról szóló gyerekművet?
- Az elmúlt néhány évben milyen mozifilmeket láttatok a mesterséges intelligenciáról?
- Miközben néztétek a filmet, vagy utána, milyen érzések, gondolatok fogalmazódtak meg bennetek? miért?

Főkérdések:

- Mit jelent számotokra a mesterséges intelligencia?
- Saját szavaiddal megfogalmazva milyen definíciót tudnál adni a mesterséges intelligenciáról?
- Mit gondoltok, mi a különbség a robotok és a mesterséges intelligencia között?
- Eddig milyen területeken találkoztatok a mesterséges intelligenciával a valós életben?
- Ezekon a területeken a mesterséges intelligencia megjelenésével, illetve az általa nyújtott szolgáltatásokkal elégedett voltál?
- Ha nem, mit hiányoltál? ha igen, mi tetszett leginkább?
- Mit gondoltok, mely területeken lehet számítani a mesterséges intelligencia gyors megjelenésére?
- Ez milyen veszélyeket és lehetőségeket rejt magában?
- Szerinted Te fel vagy készülve a veszélyek mérséklésére?
- Ha igen, akkor mit javasolnál másoknak, mit tegyenek?

- Ha ezt a kerekasztal beszélgetést 10 év múlva, 2032-ben ismételnék meg, mit gondolsz, mennyivel lenne másabb a világ a mesterséges intelligenciának köszönhetően?
- Boldogabban, vagy gondterheltebbek lennénk?
- Mely problémákra fog adni megoldás a mesterséges intelligencia az elkövetkező években?
- Egyre hangsúlyosabban fogalmazódik meg az igény, hogy a mesterséges intelligenciát csak addig és csak olyan irányokba szabad fejleszteni, hogy az ne veszélyeztesse az emberiség fejlődését. Mit gondolsz, ez hogyan lehetséges?
- El tudsz képzelni olyan helyzetet/területet, amikor a mesterséges intelligencia nem gondolkodik humánusan?
- Mit gondolsz, a Te szakterületeden a mesterséges intelligencia milyen változásokat fog hozni?
- Ezek jók, vagy inkább rosszak lesznek a számodra?

Zárókör:

- Ha a mai beszélgetést egyetlen mondattal kellene zárnod, mi lenne a tanúsága a számodra?

A kutatás eredménye

A fókuszcsoport tagjai ismerték egymást így a beszélgetés könnyen elindult a bevezető kérdéssel. Hasonlóan a senior szakértőkhöz, gyermekkorukban mindannyian találkoztak olyan filmekkel, amelyekben mesterséges intelligencia vagy robotok szerepeltek és azóta is rendszeresen néznek olyan filmeket, sorozatokat, amelyek valamilyen formában érintik ezeket. Első gyermekkori emlékeiket a témában az általános iskolai időszakhoz kötötték: főleg filmeket soroltak fel, de a Lego játékokat is többen említették. A meghatározó filmélmények között szerepeltek a következők: Én a robot, Robotok, Terminátor, Mátrix, Én vagyok anya, Johnny 5, Bosszúállók, Transformers, Knight Rider, Gungame, Starwars, 2001: Űrodüsszeia, Wall-e, Úrdongó, Big bug, Ex machina. Videójátékokat is megneveztek, többek között a Mórió és a Hero neve hangzott el.

A mesterséges intelligenciáról szóló történeteket már gyermekként is ambivalens érzésekkel élték meg: milyen mértékben humánus („barát vagy ellenség” pl. a Terminátor filmek), etikus (különbséget tesz-e a jó és rossz között). Egyetértettek abban, hogy a filmek általában negatív színben tüntetik fel a mesterséges intelligenciát, ez pedig torz képet alakít ki az emberek többségében, akik pusztán ezek alapján alkotnak fogalmat a mesterséges intelligenciáról. Ezen tapasztalatok nyomán a résztvevők többsége inkább disztópikus jövőképet vizionált. A szakértői csoporttal szemben, ahol a filmek az inspirációk és komplex problémahelyzetek megoldási lehetőségeinek forrásaként jelent meg, a junioroknál az emocionális tartalmak (fenyegetés) domborodtak ki. Ennek hátterében állhat a személyes érintettségük, azaz az eljövendő életükben egyre nagyobb teret kap és meghatározó lesz a mesterséges intelligencia jelenléte.

A junior szakértők emellett fontosnak tartották azt is, hogy a mesterséges intelligenciával kapcsolatos társadalmi szinten kialakuló sztereotípiák alakításában a filmes ábrázolások meghatározók, ahogyan egy személy a mesterséges intelligenciával, robotokkal kapcsolatos valóságot észleli, megítéli saját életét, helyzetét, kialakítja a reakcióit. Ezek között is külön kiemelték az érzelmi befolyásolást, az ismeretlentől való félelem is alapja lehet

annak, hogy túlhangsúlyozzák a robotok szerepét a filmekben. Az egyik résztvevő szerint a filmkészítők a leendő veszélyekre hívják fel a figyelmet, ami alapvetően akár pozitív is lehetne, mert motiválhatja a tudatos felkészülést. Ezzel szemben úgy vélik, a hangsúly túlságosan eltolódott a disztópia irányába, amely nem az aktív megküzdésre készít fel, hanem kiszolgáltatottságot vetíti előre.

A következő kérdés arra irányult, hogy van-e olyan filmélményük, melyben a robotok, illetve a mesterséges intelligencia pozitív színben tűnik fel. A résztvevők több filmet is felidéztek, mint például a Transformers, a Knight Rider (bár ennél a filmnél a megítélés inkább vegyes), illetve olyan filmtartalmakat, melyekben a család olyan robotokat használ, amelyek megkönnyítik a család életét. A filmekkel foglalkozó ráhangoló kérdések alapján kibontakozott egy konszenzuson alapuló vélemény, amely a senioroknál is hasonlóan alakult: a filmek többségében a mesterséges intelligencia, illetve a robotok negatív szerepet játszanak, ahol mégis pozitív kép alakul ki róluk, ott rendszerint a robotnak az embert segítő archetípusa kerül előtérbe.

Ahogy a senior szakértőknél, úgy a junioroknál is a mesterséges intelligencia és a robot fogalmának megkülönböztetése volt a következő téma. Egyetértettek abban, hogy a robot fizikai testet öltött mesterséges intelligenciaként is definiálható. Többek szerint fontos szempont, hogy a robotot milyen módon programozzák, mivel a mesterséges intelligencia képes tanulni, addig egy egyszerű robot csak egy meghatározott utasításkészlet szerint képes cselekedni. Míg a robot egy egységet képez (fizikai teste van), addig a mesterséges intelligencia egyszerre több, vagy sok helyen is megjelenhet, különböző formákban, illetve fizikai testekben. A mesterséges intelligencia e megközelítés szerint inkább adathalmaz, felhő, ami nincs fizikai helyhez kötve. Egy újabb megközelítés az érzelmek jelentőségére fókuszált: eszerint a robotok szentelen gépek, míg a mesterséges intelligencia (amelyik akár egy robotban is megjelenhet) sokkal fejlettebb, a kommunikáció, akár az emberekkel történő interakció tekintetében, lényegében rendelkezik alapszintű empátiával. Egyik résztvevő így fogalmazott: „Mivel a mesterséges intelligencia tanul, ezért elképzelhető, hogy idővel képessé válik arra, hogy az emberért feláldozza magát, mert az érzelmek is megjelennek a tanulási folyamat következtében.”

A mesterséges intelligencia definíciójának saját szavaikkal való megfogalmazása gondolkodásra készítette a csoporttagokat, többféle módon közelítették meg: „olyan program, ami arra lett kifejlesztve, hogy tanuljon, fejlődjön azért, hogy a mi életünket könnyebbé tegye”, „pozitív változást hoz az életünkbe, de meg kell ismernünk ahhoz, hogy ne féljünk tőle”, „se nem jó, se nem rossz, hanem az alkotójának egy tükörképe”, „ha képes a tanulásra, akkor képes a változásra, tehát nem biztos, hogy olyan lesz, mint amilyennek eleve megteremtették”. Lényegében egy, az embert szolgáló társ/teremtény, amelynek a fejlődése feltétlenül kontrollt meg kell őrizni. A senior szakértők a definícióban főleg a technikai paramétereket vették sorra. Egyik csoportban sem volt megegyezés egy jól körülhatárolható fogalmi megközelítésről.

Fontos szempontnak tartották azt, hogy milyen generációnak készül a definíció, mivel a résztvevők szüleinek sokkal nehezebb elfogadni ezeket a fogalmakat, mint a saját leendő gyerekeiknek, akik már bele fognak születni ebbe a közegbe.

A junior fókuszcsoport résztvevői a seniorokhoz hasonlóan a mesterséges intelligenciával az élet minden területén találkoztak: az online kereskedelmi platformok, az online

közösségi média, a biztonsági berendezések és megoldások, a személyi asszisztensek, a metaverzumok. Saját szakirányukban, az információbiztonság területén, a mesterséges intelligencia az arcfelismerésben, a rendszámfelismerésben, a biometrikus azonosításban, mozgáskövetésben, a hangazonosításban, a kézírás felismerésében jelent meg.

Az okostelefon használata természetes velejárója a junior szakértők huszonéves korcsoportjának, így annak lehetőségeit és a rajta futó alkalmazásokat jól ismerik. Az ujlyomat felismerés, a beépített személyi asszisztens, a tanulást (matematikát) segítő alkalmazás, a hangos szöveg automatikus gépi leírása, a gyerekekkel beszélgető virtuális dada szerepeltek a mesterséges intelligenciával támogatott említések között. Említették, hogy bizonyos alkalmazások sajnálatos módon az országhatárolás miatt nem, vagy csak kiskapukkal érhetőek el, illetve az alkalmazások nyelve rendszerint valamelyik világnyelven érhető csak el ezért, ahhoz, hogy ezeket érdemben használni lehessen, beszélni kell a magyar mellett valamilyen világnyelven is, például angolul. A matematikai megoldást segítő alkalmazással kapcsolatban megkérdeztük, hogy a résztvevők mennyire bíznak meg a mesterséges intelligenciában, mint szolgáltatásban. Általában vegyes a megítélésük, s inkább a saját számításaik ellenőrzésére használták, mivel számos esetben rossz megoldást adott az alkalmazás. Amikor a hallgató által számított eredmény és a gép által számított eredmény megegyezett, akkor a hallgató elfogadta, amikor viszont nem, akkor a hallgató a saját eredményét fogadta el. Ahogy egyikük megfogalmazta: „a gépi intelligenciának csak akkor hiszek, amikor a humán intelligenciával azonos eredményt ad”.

Felvetettük a résztvevőknek, hogy véleményünk szerint a mesterséges intelligencia bizonyos területeken az átlaghoz képest jobban fog fejlődni, s lesznek olyan területek, amelyek lemaradnak. A csoporttagok úgy látják, hogy önmagában nem csak a különböző területek, hanem a különböző országok között is komoly fejlődésbeli különbségek lesznek. Az egészségügyi területen járatos csoporttag a húzóágazatok között az egészségügyet és ezen belül az orvost mellőző diagnózis felállítást említette, ehhez csatlakoztak a többiek az egészségügyi rendszereket ellátó és kiszolgáló gépek fejlődésével: (pl. automatikus szállítógépek). A csoporttagok a komplex biztonsági megoldásokat, a balesetek, káresemények megelőzésének a lehetőségét, az intelligens otthonok (domotika) széles körben történő elterjedését, az információbiztonság területén a logfájlok eddiginél lényegesen fejlettebb és gyorsabb elemzését, a kibervédelmet és ennek másik oldalát a kibertámadásokat, az oktatást és ezen belül a személyre szabott oktatást sorolták még fel.

A következő kérdéscsoport a mesterséges intelligencia veszélyeinek és lehetőségeinek a feltérképezésére irányult. A veszélyek között leghangsúlyosabban, ahogyan a seniorok csoportjában is, a makroszintű munkaerő-piaci változások, a munkahelyek elvesztése szerepelt, illetve ezzel kapcsolatban a hozzátették a proaktivitás fontosságát, azaz az embereknek fel kell készülniük még időben ezekre a változásokra, mert ha erre nem képesek, akkor vesztesei lesznek a mesterséges intelligencia elterjedésének. Veszélynek érzik még a mesterséges intelligencia önállóságának fokozódását, ami révén öntudatra ébred, s „elszabadul”. Ennek elkerülésére fontosnak tartják a megfelelő „vészleállítógomb” beiktatását. A lehetőségek között, központi szerepet kap az ember kényelmének és biztonságának fokozása, ahogy a beszélgetés során ez több esetben is elhangzott.

A junior szakértők úgy gondolják, hogy a mesterséges intelligenciával rendelkező eszközök, berendezések, gépek (pl.: gépkocsi) „egymáshoz kapcsolódni tudjanak azért, hogy közösen tudjanak jó döntéseket hozni, ehhez pedig megfelelő infrastruktúra (pl.: 5G

hálózat) szükséges”. A mesterséges intelligenciák összekapcsolását egyébként ennek ellenére vegyesen ítélik meg, utalva arra, hogy a mesterséges intelligencia együttműködése révén olyan irányba fejlődik, ami lehet rossz (a gépek az ember ellen fordulnak) és lehet jó (az összekapcsolódás a biztonságot fokozza) is. Ezek a tényezők a kontroll és a biztonság szükségletei köré szerveződnek.

A beszélgetés során a generációs különbségek visszatérő szempontként vetődtek fel, mert a mesterséges intelligenciát a különböző korosztályok eltérően ítélik meg. A fókuszcsoporthoz tagjai szerint saját generációjuk – különösen azok, akik nem rendelkeznek megfelelő műszaki, informatikai ismeretekkel – gyakran nem is tudják, hogy mit jelent a mesterséges intelligencia, vagy ha tudnak valamit róla (pl. a filmélmények alapján), nem is érdeklődnek iránta, vagy negatív, elutasító álláspontra helyezkednek, vagy nem képesek felismerni, hogy az adott területen mesterséges intelligencia dolgozik. A fiatalok közül kevesen vannak, akik időt és energiát fordítanak arra, hogy szakmai-tudományos források alapján alakítsák ki saját véleményüket a mesterséges intelligenciáról. A csoport véleménye alapján a korábban megfogalmazott elvárt proaktivitás hiányzik a fiatalokból, amely a társadalmi szintű megküzdés kulcsa lenne. A résztvevők – bár személyes szabadságukat fontosnak tartják – elfogadják, hogy a mesterséges intelligenciára épülő alkalmazások folyamatosan megfigyelik őket, „nem tudunk mit tenni ellene”.

A fókuszcsoporthoz tagok szüleinek generációja – 45-60 év közötti emberek – a fiatalokhoz képest még kevésbé értik meg a mesterséges intelligenciát, félnak tőle, féltik tőle a jövőjüket, a munkahelyüket. A középkorú emberek még inkább a filmélményekre, a bulvársajtó (ál)híreire hagyatkoznak, ami összességében egy nagyon negatív, elutasító attitűdöt eredményez.

A résztvevőkhöz képest 10-15 évvel fiatalabb generáció (kistestvér, keresztgyerek, rokongyerek) és a résztvevők között is meglehetősen nagy különbség fogalmazható meg. Ez a legfiatalabb generáció már a mesterséges intelligencia jelenlétébe született bele, természetes számára, hogy előbb tanulja meg az okostelefon, tablet használatát, mint hogy beszélni tudjon. A mesterséges intelligenciával támogatott eszközök használata a kisgyerekek számára természetes, így az ezeknek az eszközöknek a használata révén kialakuló képességeik és készségeik már jobbak, fejlettebbek, mint a huszoneveseké. A mesterséges intelligenciával támogatott eszközök és tartalmak annyira izgalmasak a kisgyerekek számára, hogy szinte egész nap azzal foglalkoznak, így szocializációjukban ezek az eszközök és tartalmak egyre nagyobb szerepet töltenek be. Természetesen ez nem igaz mindenre – a résztvevők is hoztak fel ellenpéldát – de magukhoz képest úgy gondolják, hogy a mesterséges intelligencia sokkal dominánsabban volt és van jelen a kisgyerekek életében és fejlődésükben meghatározóbb hozzájuk képest: „míg nekünk meg kellett tanulni a mesterséges intelligencia használatát, addig ők ebbe már beleszülettek, számukra ez már természetes”.

A következő beszédtema a médiában és a filmekben is a valódi világ mellett kialakuló mesterséges világok (virtuális világ, metaverzum) megjelenése volt, amelyek informatikai alapját a mesterséges intelligencia jelenti. Ezzel kapcsolatban a fókuszcsoporthoz tagjai úgy gondolják, hogy jelenleg az emberek többsége még különbséget tud tenni a valós és a virtuális világ között, ugyanakkor az elkövetkező évtizedekben számolni kell azzal, hogy a *Mátrix* című filmben látottakhoz hasonló világ jön létre, azaz egyre inkább elmosódik a különbség a fizikai valóság és a virtuális tapasztalatok között, s az embernek egyre nehe-

zebb lesz tudatosítani az élmények forrását. A virtuális világok és a metaverzumok megváltoztatják emberi kapcsolatainkat, az interperszonális kapcsolatok átalakulnak, egyre nagyobb részük már nem a fizikai világban realizálódik, ez a téma ellentétes nézőpontokat hozott a felszínre. Az egyik résztvevő szerint a szenzortechnika és aktuátortechnika annyit fog fejlődni, hogy egyre kevésbé fog feltűnni a fizikai kontaktus hiánya. Egy másik résztvevő szerint ezek a mesterséges intelligencia által létrejött virtuális világok segíthetnek majd abban, hogy az introvertált, a fizikai világban félénk és visszahúzódó emberek megnyíljanak. A harmadik résztvevő szerint nem pótolható a „kézzel fogható”, valós emberi érintés, jelenlét semmilyen eszközzel, legfeljebb részben helyettesítheti azt.

A fókuszcsoport résztvevőit arra kértük, hogy próbálják „megjósolni” a jövőt, vagyis azt, hogy mi lesz 10 év múlva, 2032-ben a világban. Az egyik csoporttag Elon Musk nevét és tevékenységét meghatározónak tartja ebben a folyamatban, megemlítve a testbe építhető chipet, amelyik fogyatékos, lebénuult embertársaink számára visszaadhatja a mozgás szabadságát. Egyetértettek abban, hogy a technológia egyre gyorsabban fejlődik, alakulhat jó és rossz irányban is és számos útválasztási döntés lehetséges, ami megnehezíti a reális tényeken alapuló jövőbelátást. Egyikük így fogalmazott: „lehet, hogy 10 év múlva a mesterséges intelligencia kutatás ott fog állni, mint ma, de lehet, hogy 10 év múlva a mesterséges intelligencia olyan szinten lesz, hogy mindenki önvezető autókkal fog járni”. Egy másik vélemény szerint „én sokkal nyugodtabb lennék akkor, ha egy pozitív világgép lenne 10 év múlva és azt mondhatnám, hogy a mesterséges intelligenciának köszönhetően az én szüleim biztonságban vannak az intelligens otthonukban”. Bár 10 év múlva a fókuszcsoport tagjai alapvetően egy boldogabb, biztonságosabb, kényelmesebb világban szeretnének élni a mesterséges intelligenciának köszönhetően, azt azért nem szeretnék, ha annyira elkényelmesednének, hogy az élet minden területén már csak az unalmas élet „élménye” maradna. Fontos kérdés volt a 10 évvel későbbi világ leírásakor, hogy mennyire lehet majd átadni a teljes irányítást a mesterséges intelligenciának. Az önvezető gépjárművek technikai megoldásait elemezve ismét megerősítették a résztvevők, hogy az egyes gépjárművek egymáshoz és a nagyobb rendszerekhez (pl.: intelligens város) kapcsolódása inkább csökkenti a biztonsági kockázatokat, de az nem lehet jó fejlesztési irány, hogy amikor a mesterséges intelligencia az autóban nem tud döntést hozni, akkor hirtelen átadja az irányítást az embernek, aki lehet, hogy éppen alszik az önvezető autóban. A hangsúly azon van, hogy „amikor az ember megbízik a mesterséges intelligencia döntéseiben, különösen a gyors döntéseiben, akkor nem jó, ha hirtelen át kell venni az irányítást, mivel az ember reakcióideje lassúbb, mint a mesterséges intelligenciáé, ezért szinte biztos, hogy rossz döntést fog hozni, vagy ösztönszerűen rosszul fog cselekedni”. Ugyanakkor, amikor nincsenek ilyen helyzetek, akkor az elvárás az, hogy a mesterséges intelligencia biztosítja a biztonságot, s ha a kontroll biztonságosan átvehető, akkor azt az ember bármikor megtehesse. A fókuszcsoport tagjai a kontrollt is szükségesnek tartják, az automatikusan működő rendszereknél fontosak az olyan biztonsági megoldások, amikor meghibásodás, vagy kibertámadás esetén nem működik, vagy rosszul, vagy veszélyesen kezd el üzemelni a gép, akkor legyen egy vészleállító „gomb”, aminek aktiválása után az eszköz biztonságosan leáll. Az optimális megoldást a résztvevők abban látják, hogy a technika fejlődésének köszönhetően növekszik a bizalmunk a mesterséges intelligenciával támogatott eszközök irányába, ugyanakkor ez a bizalom csak addig mehet el, amíg még lehetőség van bármikor biztonságosan átvenni az irányítást a mesterséges intelligencia felett.

A jövő globális problémáira, vagy legalábbis azok egy részére a mesterséges intelligencia a résztvevők szerint képes lesz valamilyen megoldást kínálni. Ugyanakkor nem lenne szabad az életünket csak a mesterséges intelligencia megoldásaira bízni, mivel a globális problémák megoldásánál most és a jövőben egyaránt szükség van olyan felelősségteljesen meghozott emberi döntésekre, amelyek aztán különböző törvényekben, szabályokban, szabályzatokban tudnak realizálódni. Ez természetesen nem zárja ki a lehetőségét annak – vélik a résztvevők – hogy a döntéshozók ne támaszkodjanak a mesterséges intelligencia döntéstámogatására. Épp ellenkezőleg: „a humán döntés jobb lehet, ha a sok adatot a mesterséges intelligencia dolgozza fel, hogy az ember jobban átlássa az adatokból levonható következtetéseket” – fogalmazott az egyik résztvevő.

Az oktatás során mi gyakran említjük a Sheridan-féle skálát, ami a gép autonómiájával foglalkozik. A fókuszcsoport tagjai úgy gondolják, hogy az elképzelhető, hogy a jövőben a gépek tanítják/fejlesztik a gépeket, s így időről-időre egyre fejlettebbé válhatnak, ugyanakkor a gépeket tanító gépeknél olyan alapértékek tanítása kötelező, amelyek alapvetően az asimovi törvényekre vezethetők vissza. A képet árnyalja, hogy a mesterséges intelligenciában rejlő erőt a politikai hatalom saját céljaira is felhasználhatja, ezért fontos, hogy egy független tudományos-szakmai tanácsadó testület közösen határozza meg a fejlesztési irányokat, melyek alapvetően az emberiség érdekeit szolgálják. A tanítással kapcsolatban a fókuszcsoport tagjai között diszkusszió alakult ki arról, hogy mit kellene megtanítani a mesterséges intelligenciának: (1) tanulja meg az emberiség egész történelmét, majd ennek alapján hozzon döntést, vagy (2) az említett tanácsadó testület válogassa ki, hogy mi az a tudás, amit a mesterséges intelligenciának meg kell tanulnia ahhoz, hogy humánus döntéseket hozzon, vagy (3) a mesterséges intelligenciával a jót és a rosszat egyaránt meg kell tanítani, de azt is, hogy meg tudja különböztetni a jót a rossztól. A vitát azzal a konszenzussal zártuk, hogy ha a résztvevők majd, mint végzett mérnökök a mesterséges intelligencia tanításával, fejlesztésével foglalkoznak, akkor fontos, hogy a mesterséges intelligencia mindent tanuljon meg az emberi történelemből, hogy megfelelően megalapozott döntést tudjon hozni, de fontos, hogy megtanítsuk a mesterséges intelligenciát arra is, hogy mi a jó és mi a rossz, tehát meg tudja különböztetni ezeket a fogalmakat és az ezekhez kapcsolódó konkrét történéseket egymástól.

A résztvevők markáns különbséget tettek a mesterséges intelligencia polgári és katonai fejlesztési lehetőségei és irányai között. A polgári fejlesztésekkel kapcsolatban úgy gondolták, hogy fontos a fentebb említett nemzetközi, független, az emberiség fejlődését szem előtt tartó, humanista szakértőkből álló mesterséges intelligenciát tanító csapat. Fontos az is, hogy a csapat tagjai különböző szakterületekről érkezzenek, s pályafutásukkal szolgáljanak rá arra, hogy a mesterséges intelligencia tanítása során a humanista gondolkodás vezérli őket.

Egyik résztvevő felvetette annak gondolatát, hogy ahhoz, hogy a mesterséges intelligenciát hatékonyan és humánus elvek alapján lehessen fejleszteni, szükség van arra, hogy még jobban megismerjük azt, hogy az emberek hogyan működnek, hogyan gondolkodnak, hogyan, és milyen értékek mentén hoznak döntést. Az emberi gondolkodás (és álmódás) képekként történő megjelenése – ahogy az egyik résztvevő felidézte tudományos érdeklődése egyik példáját – a mesterséges intelligencia segítségével nem csak az emberi gondolkodás még jobb megértése miatt szükséges, hanem azért is, hogy ezáltal jobb mesterséges

intelligenciát lehessen fejleszteni. A résztvevők alapvetően úgy gondolják, hogy a mesterséges intelligenciát valamiféle érzelmi intelligenciával is fel kellene „okosítani”: „ha megtanítod neki, hogy mi a boldogság, akkor meg kell neki tanítani, hogy mi a szomorúság is, hogy a mesterséges intelligencia különbséget tudjon tenni a két fogalom között”. A mesterséges intelligenciát nem csak ellenőrzött módon tanítjuk, hanem abból is tanul, amit a külvilágból érzékel, tehát fel kell készíteni arra is, hogy a külvilágból szerzett információkat az emberiség számára és érdekében megfelelő módon tudja majd feldolgozni.

A fókuszcsoport tagjai számára – mivel mérnöki tanulmányokat folytatnak – nem voltak ismeretlenek a lágyszámítási módszerek, így a fuzzy logika sem. Ennek mentén úgy gondolják, hogy az igen-nem, jó-rossz bináris logikát érdemes jobban árnyalni a mesterséges intelligencia tanításánál. „A jó irány az lenne, ha a mesterséges intelligencia humánusan gondolkodna”, s a jövőben a felhasználói területek nagy többségénél agressziómentesen viselkedne, vagyis „nem bántaná az embert, nem korlátozná az ember szabadságát”. Az álláspontot diverzifikálja, hogy az alapvetően békésen viselkedő mesterséges intelligencia is kerülhet olyan helyzetbe, amikor agresszívan kell viselkednie. Az egyik résztvevő a családi humanoid robotot hozta fel példaként, aki alapvetően kedves, segíti a családot, részt vesz a gyerek nevelésében, de amikor egy betörő támadja meg a családot, akkor akár agresszió árán is megvédi a betörővel szemben a családtagokat. Itt az emberhez hasonlóan a támadás elhárításának a mértéke lehet a megoldás. A robotnak képesnek kell lennie felismerni az adott helyzetet, s ennek alapján csak a szükséges mértékig szabad ellenlépést tennie a támadóval szemben. Ez – véli az egyik résztvevő – az alapján is lehetséges, hogy „a mesterséges intelligencia a szenzorok segítségével folyamatosan monitorozza az ember biológiai állapotát, s ha a védendő személy pulzusa felgyorsul például a stressz miatt, akkor tudatosul a robotban, hogy baj van, cselekednie kell”.

A bizalom kérdése felmerült a beszélgetés során a társas intelligenciával rendelkező robotokkal kapcsolatban is a robot/humanoid és az ember dada, illetve nagymama vonatkozásában. A robot sokkal tudatosabban fog viselkedni, ami kiszámíthatóbb (tehát egy humán dadával szemben akár megbízhatóbb is lehet), ugyanakkor fontos, hogy a humanoid képes legyen a családban, a családi életciklusok változásával együtt fejlődni, alkalmazkodni. Különösen érdekes a speciális nevelési igényű gyerekek mellett a jövőben megjelenő humanoid dada, amelyik az átlagos szülőkhöz képest sokkal nagyobb tudással rendelkezhet az ilyen gyerekek nevelésében, fejlesztésében, tehát – elvileg – sokkal eredményesebben képes az ilyen gyerekekből kihozni azt a maximumot, amivel később minél teljesebb életet tudnak majd élni.

Mivel a résztvevők biztonságtechnikai mérnök szakos hallgatók, ezért külön kérdésként foglalkoztunk a katonai robotok használatával. A katonai robotok esetében úgy gondolják, hogy a mesterséges intelligencia képes olyan feladatokat ellátni, amelyek révén a katonai akciók eredményesebben hajthatók végre, kevesebb a saját oldalon a humán katonai veszteség, a katonák biztonságos távolságból tudják nyomon követni, vagy irányítani a robotkatonákat. Ugyanakkor veszélyként merülhet fel, hogy ha a robotkatonára felismeri, hogy alkalmazásának célja az, hogy feláldozza magát az emberért, akkor meglehet, hogy fellázad ez ellen. A megoldást az egyik résztvevő abban látja, hogy „mekkora szabadságot engedek a robotkatonának. Ha a robotkatonára szabadsága erősen beszabályozott, akkor a katonát képes kiszolgálni, támogatni, például azzal, hogy viszi a nehéz csomagokat, vagy képes felderíteni és akár megsemmisíteni az ellenséget, de azt nem tudom elképzelni, hogy

mellettem egyenrangú félként harcoljon”. Miközben több szakirodalom foglalkozik a robot-robot harccal, vagyis amikor a robotkatonára harcol robotkatonára ellen, a fókuszcsoporthoz tartozó tagjai úgy vélik, hogy ez nem minden esetben jelent megoldást. Három okot neveztek meg ezzel kapcsolatban: (1) ha a robotot a háttérből ember irányítja, és az irányítás megszűnik, akkor a robot harcképtelenné válik, tehát könnyen megsemmisíthető az ellenség részéről, (2) az elektromágneses impulzusfegyverek képesek akár az önjáró, akár a háttérből vezérelt robotkatonákat harcképtelenné tenni, (3) az egymással szemben álló robotkatonák technikai fejlettségbeli különbsége eldöntheti a harc kimenetelét. Megemlítették azt is, hogy miközben a hagyományos hadviselésnél alkalmazott szabályok (rules of engagement /ROE/) viszonylag egyértelműen fogalmazzák a harci cselekményekkel kapcsolatban, addig különösen az autonóm harcjárművek, illetve humanoid robotkatonák esetében ezek a direktívák még nem állnak rendelkezésre.

A mesterséges intelligencia katonai felhasználási lehetőségeivel kapcsolatban ismét felmerült az a kérdés, hogy mekkora autonómiát engednének meg a résztvevők a jövő katonai és polgári humanoid, vagy általánosságban mesterséges intelligenciával támogatott robotjai számára. A kérdés elbizonytalanította a résztvevőket, részint amellyel próbáltak meg állást foglalni, hogy a gépeknél inkább korlátozni és minimalizálni kell az agressziót, illetve, ha a robot lecsatlakozik a hálózatról, akkor át kell váltania egy egyszerűbb működésre, ami minimális funkciókat engedélyez a számára (pl., hogy megkeresse a legközelebbi felcsatlakozási pontot), bár katonai robotoknál erre nem láttak reális megoldást, pont a speciális felhasználás miatt.

A résztvevők 10 év múlva, már végzett biztonságtechnikai mérnökként úgy gondolják, hogy olyan szakmai területen helyezkednek el, ahol napi szinten fognak találkozni a mesterséges intelligencia lehetőségeivel, így például az információbiztonság, a haditechnika, vagy a nemzetbiztonság területén. Örömmel vennének részt a mesterséges intelligencia fejlesztésében, tanításában.

A fókuszcsoporthoz tartozó beszélgetés zárásaként a következő összegző mondatokat fogalmazták meg a résztvevők: „érdekes beszélgetés volt, jó volt megtapasztalni a sokféle, akár egymástól eltérő véleményt is”, „számomra is nagyon jó volt, az eltérő véleményünk alapja szerintem az volt, hogy ki mennyire bíz meg a mesterséges intelligenciában”, „nagyon hasznos, érdekes, informatív beszélgetés volt”, „nagyon szerteágazó a véleményünk, s igazából nem tudjuk, hogy mi lesz az elkövetkező tíz évben a mesterséges intelligenciával, csak sejtéseink vannak”, „a sokszínű gondolkodásmód és bizonytalan jövőkép mellett meg kell találnunk egyfajta konszenzust, ami révén kialakítható a mesterséges intelligencia és a közös jövőnk”.

A JUNIOR SZAKÉRTŐK VÉLEMÉNYE ALAPJÁN MEGFOGALMAZOTT KÖVETKEZTETÉSEK, JAVASLATOK

A junior szakértők a mesterséges intelligencia területén járatos egyetemi hallgatók voltak, akik életkoruknál fogva a senior szakértőktől eltérő tapasztalatokkal és idői perspektívával tekintettek a kérdésekre. A mesterséges intelligenciával kapcsolatos maradandó gyermekkori film és egyéb élményeik emocionális hatása élénken élt bennük. A junior szakértők egyetértettek abban, hogy a mesterséges intelligencia fejlődési tempójának és irányának kiszámíthatatlansága következtében a potenciális jövőképek széles határok között mozognak. Úgy vélték, hogy ha a társadalmi tudatosság és proaktivitás helyett az érdeklődés

hiánya és a félelmek alakítják az emberek attitűdjeit, ezzel a passzivitással a technopesszimizista forgatókönyv önbeteljesítő jóslatként valósulhat meg. A disztópikus jövőkép visszaterő eleme volt a megnyilvánulásainak, amely reális fenyegetés lehet az életük alakulásában. A mesterséges intelligencia társadalmi megítélésében egyhangúan megegyeztek abban, hogy a középkorúak és idősek korosztályában a félelem dominál, amely egyrészt a média-reprezentációra vezethető vissza, másrészt a bizonytalanság következtében kialakuló természetes emberi reakció. A fiatal felnőttek, azaz a saját generációjuk nem ismeri fel a mesterséges intelligenciát, tájékozatlanok, nem is foglalkoztatja őket, ezért a veszélyeket sem méri fel, kiszolgáltatottak, beletörődő attitűddel viszonyulnak a mesterséges intelligenciához. A gyerekek széles körében már szocializációjuk természetes velejárója a mesterséges intelligencia jelenléte és számukra nem jelent fenyegetést pl. az adataik felhasználása, amit az idősebbek kevésbé fogadnak el. A generációk attitűdjeinek eltérései szintén visszaterő gondolati szervezőerő volt.

A mesterséges intelligencia és a robotok definiálása a junior szakértőknél nem eredményezett egységes álláspontot, ahogyan az elméletalkotók is különböző meghatározásokat használnak, inkább gyűjtőfogalomként értelmezhető. A mesterséges intelligencia feletti kontroll szükségletét hangsúlyozták, mint elengedhetetlen komponens, pl. egy biztonságot nyújtó vészleállító gomb formájában, amely a fenyegetéssel szembeni aktív megküzdés eszközeként értelmezhető és a vágyott optimista forgatókönyvhöz vezet. A mesterséges intelligencia veszélyei közé sorolták a munkanélküliség tömegessé válását, illetve a technika fejlődésével egyre fokozódó autonómia következtében kialakuló „öntudatra ébredést” és ezzel társultan az emberi faj elleni agresszív fellépést. A mesterséges intelligencia mindkét esetben az emberiség szolgálata helyett az alárendelt szerepből riválissá válik, ami manapság a félelemmel együtt járó bizalmi deficitet erősíti.

A junior szakértők a mesterséges intelligencia alkalmazásainak széles körét ismerik [1], maguk is használnak néhányat leginkább az okostelefonok segítségével. Ezek hátrányaként a nyelvi és a földrajzi elérhetőség szabta korlátok mellett nagyobb szerepet kapott a megbízhatóság kérdése, példaként egy számítási feladatokat végző, téves eredményeket is generáló alkalmazást említettek.

Kiemelték még az interperszonális kapcsolatok várható minőségi változását, amely a virtuális világok és a fizikai valóság közti határok összemosódásával létrejövő multidimenzionális élményekre épül, ennek megítélése ellentmondásos volt, a résztvevők mind pozitív, mind negatív véleményeket is megfogalmaztak. A kényelem, a biztonság és boldog élet reménye határozták meg a reményeiket, de egyben azt a veszélyt látják, hogy unalmassá, „túl kényelmessé” válik a világ. A 2032-ben elképzelt világban többek között az intelligens városok, otthonok, gépjárművek biztonságos, kiszámítható és humánus döntéseket hozó működését várják. Nagy horderejű döntések esetén továbbra is felelősségteljes emberi döntéshozókat, testületeket tartanak elfogadhatónak, akik a döntéstámogató rendszereket használják.

A junior csoportban megfogalmazott vélemények tartalmában a fentiekből azonosíthatók azok a pszichológiai szükségletek, amelyeket Ryan és Deci [2] öndetermináció elméletükben írtak le. Eszerint az embereknél három alapszükséglet különböztethető meg: (1) az autonómia, azaz a szabad döntések és cselekvések lehetősége, amelyet a „nem humánus döntéseket hozó”, vagy „öntudatra ébredő”, az emberi fajjal versengő mesterséges intelligencia fenyeget;

- (2) a kompetencia, azaz az énhatékonyság érzésének igénye, amelyet a mesterséges intelligencia feletti kontroll elvesztése, és az „elkényelmesedés” fenyeget;
- (3) a kapcsolódás szükséglete, azaz a szeretetre, tiszteletre épülő kapcsolatok kialakítása, amelyet a virtuális világok multiverzuma fenyeget.

A szakértők a további témáknál is kidomborították az emberi vonások szükségességét a mesterséges intelligencia bizonyos alkalmazásainál, pl. szociális, érzelmi intelligencia szükségessége, vagy a gondjaira bízott személy védelme érdekében tanúsított agresszió. Ez utóbbi különösen a katonai mesterséges intelligencia, ill. robotok felhasználása esetén fontos kérdés, pl. az ember-robot, illetve a robot-robot harc esetén a hagyományos hadviselés szabályainak érvényessége.

A SENIOR ÉS A JUNIOR SZAKÉRTŐK VÉLEMÉNYÉNEK KOMPARATÍV ELEMZÉSE

A senior és a junior szakértők tapasztalatbeli és felfogásbeli különbségeik ellenére számos területen hasonlóan gondolkodtak.

A mesterséges intelligenciához köthető karrier háttérében mindkét csoportban felismerhető a gyermekkori film és egyéb élmények hatása a pályaválasztásra. A stabil pályaidentitással rendelkező seniorok körében kiegészült azzal, hogy a folyamatos fejlődési lehetőség és a megélhetés biztosításával elkötelezettek a szakmai munkájuk iránt. A két csoport közti generációs különbségből fakadóan a juniorok a jövőre irányuló idői perspektívából közelítették meg a kérdéseket, míg a senioroknak tágabb rálátása van a mesterséges intelligencia múltjára, jelenére, jövőjére. Mindkét csoport számos mesterséges intelligenciára épülő alkalmazást ismer, leginkább az okos otthonok (domotika), intelligens városok, autonóm járművek, egészségügyi és gyógyszeripari alkalmazások, gazdasági pl. banki, katonai területen, mezőgazdasági, környezetvédelmi, oktatási célú felhasználás.

Annak ellenére, hogy a fókuszcsoportok tagjai kompetens szakértők voltak, a mesterséges intelligencia definíciójában nem alakult ki konszenzus egyik csoportban sem, ez egybevág azzal, hogy a szakirodalomban is számos különböző meghatározás létezik, ezért inkább gyűjtőfogalomnak tekinthető. Mindkét csoportban tartalmilag a biztonságosság és kényelmesség köré szerveződtek az asszociációk, amelyet a juniorok emocionális színezettel fogalmaztak meg (jó-rossz, félelemkeltő), a seniorok pedig a racionális (technikai, gazdasági) oldalt ragadták meg. A robotok és a mesterséges intelligencia megkülönböztetésében a robotokhoz fizikai megtestesülést társítottak, mindkét csoportban megjelentek az egyik póluson a kiber-fizikai rendszerek a másokon pedig a humanoid/animoid robotok, ez utóbbiak szociális, emocionális készségeikkel felruházva. Mindkét csoportban egyetértettek abban, hogy a robotok és a mesterséges intelligencia médiareprezentációjában jelentősen több a disztópikus ábrázolás, ahol az emberi faj létét fenyegetik, ezért a mesterséges intelligencia társadalmi percepciójában a negatív emocionális tartalmak dominálnak, elsősorban a félelem jellemző.

A fókuszcsoportok szakértői részletesen kifejtették a mesterséges intelligenciával kapcsolatos kockázatokat, kihívásokat, amelyek az alábbi fő területeket érintették:

- Társadalmi szintű bizalmatlanság: a mesterséges intelligencia gyakran olyan kontextusokban jelenik meg, amely fenyegetést jelent, visszaélések eszközévé válhat, így gyanakváshoz, elutasításhoz vezet.

- Transzparencia hiánya: komplexitásánál fogva magas szinten képzett, tapasztalt szakemberek számára is kihívást jelent a mesterséges intelligencia gyakorlati alkalmazásainak kockázatelemzése, pl. a potenciális veszélyek feltárása (hibás programozás, hibás tanítás, kibertámadás), az ezek következtében fellépő anyagi vagy más erőforrásbeli károk megbecsülése.
- Torzítások, diszkrimináció, amely leképezi az emberi döntéshozatal szisztematikus hibáit, pl. az előítéletekkel terhelt (vallási hovatartozás, nemi identitás) adatokra épülő gépi tanulás eredményeképpen alakulhat ki.
- Az információ- és adatbiztonság a szabályozatlanság és átláthatatlanság miatt egyre kevésbé garantálható, az adatok felhasználása áttekinthetatlenné válik, a privát szféra és az az emberi jogok sérülhetnek, pl. társadalmi kredit rendszere.
- A mesterséges intelligenciát tisztességtelen politikai, gazdasági hatalmi célok érdekében alkalmazzák, a demokratikus értékek és az emberi jogok sérülnek.
- Munkaerő-piaci átstrukturálódás: egyrészt az emberi munka mesterséges intelligenciára épülő technológiával való helyettesítése tömeges munkanélküliséget okozhat megfelelő átképzések hiányában. Másrészt a mesterséges intelligencia fejlesztéséhez megfelelően képzett munkaerő iránti egyre növekvő kereslet kielégítése munkaerőhiányt okozhat, ma még nem ismert új szakmák megjelenése várható. A munkaerő-piaci átrendeződés a társadalmi egyenlőtlenségek fokozódásához, a társadalom szétszakadásához vezethet.

A junior szakértők szerint a generációk eltérő módon viszonyulnak a mesterséges intelligenciához. Minél fiatalabb a személy, annál elfogadóbb és természetesebb számára, az idősebbek pedig bizalmatlanok, fenyegetőnek találják. A szakértők szerint a veszélyek megelőzését szolgáló lépéseket, intézkedéseket nem lehet halogatni, a kedvezőtlen folyamatok ma még kontroll alá vonhatók, a jövő alakítható és ebben a szakembereknek is aktív szerepet kell vállalniuk.

A szakértők bíznak abban, hogy a mesterséges intelligencia 10 év múlva az emberiség jóllétét és kényelmét fogja szolgálni, hozzájárul az életminőség javulásához, humánus és etikus módon fog működni. Ennek érdekében a társadalom minél szélesebb körét fel kell készíteni a mesterséges intelligencia várható hatásaira és használatára. Mindkét csoport szerint a virtuális világok, metaverzumok és a fizikai valóság közti határ elmosódik, elképzelhető, hogy a virtuális világba terelődik a boldogságkeresés, a hús-vér emberi kapcsolatokat az avatárok által konstruált világok válthatják fel, az érzékszervi tapasztalatokat a szenzor-technika fogja pótolni.

A mesterséges intelligencia humánközpontúságának egy speciális területe a hadviselés, ahol az agresszió és a hagyományos hadviselési szabályok alkalmazása számos kérdést vet fel. A seniorok a szabályozás oldaláról közelítették meg a kérdést: ezen a területen a mesterséges humánium bevezetése szükséges, azaz az emberi beavatkozás lehetősége mindvégig adottságként jelen legyen, az ellencsapás mértéke ne legyen jelentősen nagyobb a csapásnál. Nem lehet cél az ellenfél totális elpusztítása sem, a háborúkra vonatkozó nemzetközi jogi szabályok, előírások betartása továbbra is követelmény kell, hogy maradjon. A hadviselési szabályokat ki kell dolgozni a mesterséges intelligenciával működtetett eszközökre is. A juniorok technikai szempontból vizsgálták a kérdést: a robotkatonák alkalmazásáról megállapították, hogy kockázati tényezőként számolni kell a kontrollálhatatlanná válással, ill. a háború kimenetelét eldöntő technikai fejlettséggel. A robotok és a mesterséges

intelligencia agresszív viselkedéséről a juniorok megállapodtak abban, hogy az ember testi épségének védelme érdekében elfogadható lehet, példaként egy gyermeket a betörővel szemben védelmező családi robotot képzeltek el, amely a gyermek fiziológiás reakciói segítségével vált védelmező üzemmódra.

Mindkét csoportban kifejezték az etikus mesterséges intelligencia megteremtésének fontosságát, ennek hiányában a disztópikus jövőre vonatkozó aggodalmaikat reális fenyegetésnek érzékelik. A mesterséges intelligencia ember által alkotott „teremtmény”, amely az emberi fajjal békében együtt élő társá vagy riválissá válhat a koevolúció során. A junior szakértőknél gyakrabban jelentek meg emocionális színezetű vélemények, míg a seniorok konkrét tapasztalatokat, ismereteket osztottak meg, ezáltal tényszerűbb, racionálisabb formában fejezték ki álláspontjukat.

További tartalomelemzési szempont lehet későbbi kutatásokban – mely jelen cikk kereteit meghaladja – az ember pszichológiai szükségleteinek vizsgálata a mesterséges intelligenciára vonatkozó attitűdökben. Kiindulásként szolgálhat pl. a széles körben ismert Maslow [3] féle öt szintű szükséglet-hierarchia A fiziológiás szükségletek kielégítésének biztosítása a termelés és szolgáltatások területén, már a jelenleg is alkalmazott eljárások tovább fejlesztésével, egyre kényelmesebbé teheti az életet, de a tartós komfortérzet unalomhoz vezethet. A biztonságsszükségletek a mesterséges intelligencia kiszámíthatóságára, megbízhatóságára, a kockázatok és veszélyek minimalizálására irányulnak, pl. a „védelmező robot” is ezt a célt szolgálja. A hovatartozás és szeretet szükségletének kielégítésére már ma is léteznek szociális robotok („szerethető robot”), az emberrel empatizáló, érzelmi és társas intelligenciával rendelkező mesterséges intelligencia töltheti be ezt a szerepet. A megbecsülés szükséglete akkor elégül ki, ha az ember megőrzi az autonómiáját és alkotóképességét, ezzel mások elismerését kivívja, pl. aktívan szerepet vállal a mesterséges intelligencia fejlesztésében, vagy a szabályozó testületek munkájában, élethosszig tartó tanulásal képes a munkaerő-piaci kihívásoknak megfelelni. Végül az önmegvalósítás számára tág lehetőség nyílik nemcsak a metaverzumokban, hanem a kényelmes élettel járó több szabadidő kreatív eltöltésével.

Mindkét csoport hasznosnak találta a fókuszcsoportban a tapasztalatok és vélemények kölcsönös megismerését, a közös gondolkodást, újabb nézőpontokat ismertek meg, tudatosították a fenyegetések széles körét.

ÖSSZEGZŐ GONDOLATOK, KÖVETKEZTETÉSEK

A mesterséges intelligencia fejlődésében kritikus szakaszba lépett az emberiség, egyszerre rejti magában a lehetőséget az emberek életminőségének ugrásszerű javulására, de a disztópikus jövő is realizálódhat, ahogyan explicit módon a szakértők is kifejtették. Sinderman [4] és munkatársai a mesterséges intelligenciával kapcsolatos öt attitűdöt különböztettek meg, amelyekben visszatükröződik e két irány: bizalom a mesterséges intelligenciában, a mesterséges intelligencia hasznos lesz az emberi faj számára, ezzel szemben a mesterséges intelligencia félelmetes, fenyegető, az emberi faj pusztulásához vezet, munkanélküliséget okoz számos területen. A potenciális fenyegetések csökkentése érdekében nem elegendőek a nemzetközi szervezetek szabályozó intézkedései (pl. UNESCO [5], OECD [6], [7]), társadalmi szintű fellépés szükséges a demokratikus értékek és az emberi jogok védelmében, a közjő és jóllét növelésében, a személyes adat- és információbiztonság fokozásában. A társadalmi bizalom javításához az intelligens rendszerek etikus fejlesztésében

biztosítani kell a hatékonyságot, az átláthatóságot, a professzionalizmust, ki kell jelölni a felelősségi köröket és a visszaélések lehetőségét minimalizálni szükséges, ahogyan ezt egy 700 fős szakértői közösség is deklarálta [8]. A szakértők, Tilesch és Hatamleh [9] megállapításához hasonlóan, a humánközpontú mesterséges intelligencia létrehozásában egy globálisan elfogadott, megvalósítható mesterséges intelligencia vízió kialakítását sürgetik, amely az általános emberi értékekre épül, az eltérő érdekeket harmonizálja, egységes szabályozási keretbe foglalja, a paradigmaváltás azonban csak akkor válik teljessé, ha az emberiség a lelkiismeretességére és tudatosságra alapuló új világnézetet hoz létre, ezzel biztosítva a faj fennmaradását.

FELHASZNÁLT IRODALOM

- [1] Kollár Cs.: A mesterséges intelligencia és a kapcsolódó technológiák bemutatása a biztonságtudomány fókuszában in Kiberbiztonság/Cybersecurity, edited by Z.Rajnai Biztonságtudományi Doktori Iskola, Budapest, 2019, ISBN: 978-963-449-185-9. letöltés: <https://drkollar.hu/blog/2020/01/27/kiberbiztonsag-cybersecurity-uj-ingenyesen-elerheto-kiadvany/>
- [2] Ryan, R. M., Deci, E. L., Self-determination theory: Basic psychological needs in motivation, development, and wellness. 2017. New York: Guilford Publishing
- [3] Maslow, Abraham H. A Theory of Human Motivation (1943), Originally Published in Psychological Review, 50, 370-396. letöltés: <http://psychclassics.yorku.ca/Maslow/motivation.htm>
- [4] Sindermann, C., Sha, P., Zhou, M. et al. Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. *Künstl Intell* 35, 109–118 (2021). letöltés: <https://doi.org/10.1007/s13218-020-00689-0>
- [5] UNESCO: PRELIMINARY STUDY ON THE ETHICS OF ARTIFICIAL INTELLIGENCE (COMEST), 2019.: <https://en.unesco.org/artificial-intelligence/ethics/cases#biasedai>
- [6] OECD (2019) Artificial Intelligence in Society. letöltés: <https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm>
- [7] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. letöltés: 22/05/2019, letöltés: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- [8] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems 2019. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. 1st ed.: IEEE. letöltés: <https://ethicsinaction.ieee.org/#series>
- [9] G.A.Tilesch, o. Hatamleh: Between brains, Taking back our future in the AI age, GTPublishDrive; Publishdrive ed. edition, 2020.