

Lóránt Szabolcs<sup>1</sup>

## Mesterséges intelligencia és legitim autoritás a modern hadviselésben

Az igazságos háború elméletének újraértelmezése  
az ukrajnai háború tükrében

### Artificial Intelligence and Legitimate Authority in Modern Warfare

Reinterpreting Just War Theory in Light of the War in Ukraine

#### Absztrakt

*A tanulmány az igazságos háború elmélet legitim autoritás elvének újraértelmezését vizsgálja a mesterségesintelligencia-alapú hadviselés kontextusában. A kutatás fő kérdése, hogy a mesterséges intelligencia és az autonóm fegyverrendszerek megjelenése hogyan változtatja meg a legitim autoritás elvének gyakorlati alkalmazását, különös tekintettel az elszámoltathatóság és az átláthatóság követelményeire. Az ukrajnai háború esettanulmányain keresztül elemzi az úgynevezett feketedoboz-problémát és az emberi felügyelet (human in the loop) elv gyakorlati megvalósulását. A tanulmány amellet érvel, hogy noha a mesterségesintelligencia-alapú rendszerek gyengítik a legitim autoritás hagyományos értelmezését, ez nem feltétlenül jelenti az elv érvénytelenné válását. A tanulmány következtetése szerint az elv újraértelmezése nemcsak technológiai és jogi, hanem alapvető etikai kérdéseket is felvet, amelyek további vizsgálatot igényelnek.*

*Kulcsszavak: igazságos háború elmélete, legitim autoritás, mesterséges intelligencia, autonóm fegyverrendszerek, emberi kontroll*

<sup>1</sup> Doktori hallgató, Nemzeti Közszolgálati Egyetem Hadtudományi Doktori Iskola.

## Abstract

*This paper examines the reinterpretation of the right authority principle of just war theory in the context of artificial intelligence-based warfare. The main research question is how the emergence of AI and autonomous weapon systems changes the practical application of the principle of right authority, with a particular focus on the requirements of accountability and transparency. Through case studies of the war in Ukraine, the paper analyses the “black box” problem and the “human in the loop” principle in practice. The paper argues that although AI-based systems weaken the traditional understanding of right authority, this does not necessarily mean that the principle is invalid. The paper concludes that the reinterpretation of the principle raises not only technological and legal, but also fundamental ethical questions that require further investigation.*

*Keywords: just war theory, legitimate authority, artificial intelligence, autonomous weapons systems, human control*

## Bevezetés

„A mesterséges intelligencia az emberi intelligencia sikeres folytatása más eszközökkel”<sup>2</sup> – fogalmazta át Carl von Clausewitz porosz katonai teoretikust Luciano Floridi olasz–brit filozófus és etikus.

A mesterséges intelligencia (MI) nagymértékben átalakítja az emberi tevékenységet. Ez különösen igaz a hadviselésre, ahol – akárcsak egykor a lőpor vagy a nukleáris fegyverek – forradalmi változásokat hoz, de különleges kihívásokat is rejt magában. Ezen kihívások új megvilágításba helyezik a hadviselés etikai kérdéseit.

Az igazságos háború elmélete a háború etikai megítélésének egyik legjelentősebb elméleti keretrendszere, amelynek gyökerei az ókori filozófiáig nyúlnak vissza, jelentős változata alakult ki a középkori keresztény gondolkodásban, és máig meghatározza a nemzetközi kapcsolatokat és a hadviselés szabályait.<sup>3</sup>

Az igazságos háború elmélete két fő komponensre osztható: a háborúba lépés joga (*jus ad bellum*) és a hadviselés módja (*jus in bello*).<sup>4</sup> Az előbbi hat kritériuma – igazságos ok, legitím autoritás, helyes szándék, arányosság, végső eszköz és észszerű esély a sikerre – együttesen határozza meg a háború megindításának jogosságát. A legitím autoritás ebben a rendszerben különleges szerepet tölt be, mivel összeköti a politikai legitimációt a katonai cselekvéssel, vagyis mind a háború megindítására, mind a hadviselésre vonatkozatható.

A legújabb kori hadviselésben az egyik legjelentősebb változást az autonóm fegyverrendszerek és a mesterséges intelligencia megjelenése hozta, ami alapvető kihívások elé állítja az elmélet hagyományos elveit, különösen a legitím autoritás kritériumát. Az MI és az autonóm fegyverrendszerek megjelenése különösen az elszámoltathatóság

<sup>2</sup> FLORIDI 2023: 23.

<sup>3</sup> BODA 2023.

<sup>4</sup> WALZER 2015: 21–33.

és az átláthatóság követelményeinek érvényesíthetőségét teszi problematikussá. Olyan rendszereknél, ahol a döntéshozatali folyamat nem teljesen átlátható (feketedoboz, *black box* probléma), felmerül a kérdés: hogyan értelmezhető és alkalmazható a legitím autoritás elve ebben az új technológiai környezetben?

A tanulmány ezekre a kérdésekre keresi a választ, amihez meg kell vizsgálni:

- az elszámoltathatóság értelmezését a mesterséges intelligencia feketedoboz-problémájának tükrében,
- az emberi felügyelet (*human in the loop*) elv szerepét a legitím autoritás fenntartásában,
- a döntéshozatal átláthatóságának biztosítási lehetőségeit MI-alapú rendszerek esetén.

Elemzésünk szerint a mesterségesintelligencia-alapú autonóm fegyverrendszerek alkalmazása gyengíti a legitím autoritás elvének érvényesülését, mivel a feketedoboz-probléma miatt csökken a döntéshozatal átláthatósága és az elszámoltathatóság lehetősége, amit a *human in the loop* elv csak részben képes ellensúlyozni.

A fenti kérdések megválaszolásához először szükséges az igazságos háború elméleti keretrendszerének áttekintése, azon belül kiemelten Michael Walzer amerikai politikai filozófus megközelítése, amely a legitím autoritás morális dimenzióit hangsúlyozza. Ezt követően megvizsgáljuk a makulátlan pusztítás (*immaculate destruction*) doktrínát, amely az amerikai katonai gondolkodás sajátos termékeként a technológia és a hadviselés kapcsolatára nyújt értelmezési keretet. A Clausewitz perspektíva segít megérteni, hogy a technológiai fejlesztések miként hatnak a háború politikai dimenziójára. A mesterségesintelligencia-rendszerek jelentette kihívások és az ukrajnai esettanulmányok konkrét példákön keresztül szemléltetik a legitím autoritás elvének gyakorlati érvényesülését. Az elemzés és az elszámoltathatóság új dimenzióinak vizsgálata után az összegző gondolatok választ adnak a kutatási kérdésekre, kiemelve a legitím autoritás elvének újraértelmezési lehetőségeit az MI korában.

## Elméleti háttér

### A legitím autoritás elvének értelmezési keretei: a walzeri megközelítés

Az igazságos háború elméletben a legitím autoritás elve tradicionálisan azt követeli meg, hogy a háborús döntéseket legitím, elszámoltatható hatóság hozza meg. James Turner Johnson amerikai etikus és történész szerint a mesterségesintelligencia-alapú hadviselés kontextusában ez az elv új értelmezést nyer, mivel az MI-rendszerek nem pusztán támogató eszközök, hanem *de facto* stratégiai szereplőkké válnak.<sup>5</sup> Ez alapvetően érinti mind a legitím autoritás állami szintjét, mind pedig a parancsnoki struktúrák és az elszámoltathatóság kérdését.

<sup>5</sup> JOHNSON 2022: 262–263.

Walzer *Just and Unjust Wars* című munkájában kifejtett elmélete különösen jelentős e kihívások értelmezésében. Könyve 19. fejezetében a hatnapos háború példáján illusztrálva a szerző aláhúzza, hogy bármennyire modern és gyors a háború, a legitím autoritás nem pusztán technikai, hanem morális kategória.<sup>6</sup> Ez alapvető kérdéseket vet fel a mesterségesintelligencia-rendszerek alkalmazhatóságáról, hiszen az MI-rendszereknek nincs valódi morális érzékük vagy filozófiai értelemben vett etikai megfontolásuk.

Walzer igazságos háború elmélete a morális döntéshozatal gyakorlati megközelítését képviseli. Állítása szerint létezik egy közös morális alap, amely lehetővé teszi a háborúk egységes megítélését mind a háborúba lépés joga (*jus ad bellum*), mind a hadviselés módja (*jus in bello*) tekintetében. Noha elismeri a katolikus teológusok hozzájárulását az elmélet fejlődéséhez, megközelítése elsősorban történelmi példákon alapuló társadalmi-morális értékelés. Walzer elméletének különös erőssége, hogy a morális elveket nem elvont teóriaként, hanem a gyakorlati valóság részeként kezeli. Ahogy rámutat, a hadviselés sosem működött pusztán az erő egyenlő a joggal (*might equals right*) elve alapján, hanem mindig létezett egy gyakorlati moralitás.<sup>7</sup> Elméletének időtálló relevanciáját bizonyítja, hogy könyvének egymást követő kiadásai sikeresen alkalmazták az alapelveket olyan új kihívásokra, mint a humanitárius intervenció vagy az aszimmetrikus hadviselés. Ez a rugalmas, gyakorlati megközelítés teszi különösen alkalmassá Walzer elméletét a mesterségesintelligencia-alapú rendszerek hadviselési alkalmazásának etikai értékelésére.

Walzer könyve 5. részének (*A felelősség kérdése*) koncepciói jelentősen hozzájárulhatnak a mesterséges intelligencia és az autonóm fegyverrendszerek legitím autoritásának vizsgálatához. A szerző három kulcsfontosságú megállapítása különösen releváns.

Walzer érvelése a legitím autoritásról – miszerint az nem pusztán formális (jogi) pozíció, hanem morális felelősséggel járó szerep<sup>8</sup> – közvetlenül alkalmazható az MI-rendszerek felügyeletének kérdésére. Ahogy a szerző kifejti, a hatalom gyakorlóinak elszámoltathatósága és a döntések következményeiért való felelősségvállalás<sup>9</sup> kulcsfontosságú – ez párhuzamba állítható az MI-rendszerek fejlesztőinek és üzemeltetőinek felelősségével. A kollektív felelősség kérdése<sup>10</sup> szintén releváns a mesterséges intelligencia kontextusában, hiszen a technológia fejlesztése és alkalmazása széles körű (piaci és katonai szereplők általi) együttműködést igényel, és jelentős társadalmi hatásokkal jár. A fejezet azon megállapítása, hogy a legitím autoritásnak korlátokkal kell rendelkeznie és elszámoltathatónak kell lennie,<sup>11</sup> különösen fontos a mesterségesintelligencia-rendszerek szabályozásának és felügyeletének kialakításakor.

Walzer azon tétele, hogy a legitím autoritás nem pusztán technikai, hanem morális kategória, alapvető kihívást jelent az MI-rendszerek alkalmazásában. Mivel ezek a rendszerek nem rendelkeznek valódi morális érzékkel vagy filozófiai értelemben

<sup>6</sup> WALZER 2015: 304.

<sup>7</sup> NOLAN 2015: 125–126.

<sup>8</sup> WALZER 2015: 287–288.

<sup>9</sup> WALZER 2015: 290–291.

<sup>10</sup> WALZER 2015: 296–298.

<sup>11</sup> WALZER 2015: 291.

vett etikai megfontolásokkal, felmerül a kérdés, hogy egyáltalán gyakorolhatnak-e legitím autoritást. Walzer elméletének etikai megközelítése alapján a legitím autoritás gyakorlásához szükséges a morális felelősség vállalásának képessége és a morális döntéshozatal – olyan tulajdonságok, amelyeket az MI-rendszerek a jelenlegi formájukban nem birtokolnak. Ez azt sugallja, hogy a mesterségesintelligencia-rendszerek csak mint eszközök funkcionálhatnak a legitím autoritás gyakorlásában, de önmagukban nem lehetnek annak forrásai vagy gyakorlói. Ez különösen releváns az autonóm fegyverrendszerek alkalmazásának etikai kérdéseiben, ahol a morális döntéshozatal nem delegálható teljes mértékben algoritmikus rendszerekre.

## A makulátlan pusztítás

A *makulátlan pusztítás* doktrína a vietnámi háború utáni amerikai katonai gondolkodás sajátos terméke, amely a 2000-es évek elejére vált meghatározóvá. A koncepció központi eleme az a meggyőződés, hogy a modern technológia lehetővé teszi a háború teljes kontrollját és kiszámíthatóságát.<sup>12</sup>

A doktrína szerint a csatatér teljes átláthatósága és a precíziós fegyverek kombinációja révén a háború „tisztá” és humánus módon vívható meg. William Perry volt védelmiminiszter-helyettes már 1978-ban úgy vélte, hogy a Pentagon hamarosan képes lesz mindent látni a csatatéren, eltalálni mindent, amit lát, és megsemmisíteni mindent, amit eltalál.<sup>13</sup>

A koncepció nagy hangsúlyt helyez az információs fölényre és a hálózatközpontú hadviselésre. William Owens, az Egyesült Államok haditengerészetének nyugalmazott admirálsa szerint az információs technológia első alkalommal teszi lehetővé a parancsnokok számára a mindentudó perspektívát a csatatérről valós időben, éjjel-nappal, minden időjárási körülmény között.<sup>14</sup>

Az elmélet gyakorlati alkalmazása során az ellenséget „rendszerként” kezelik, amelynek működését precíziós csapásokkal lehet megbénítani. Ez a megközelítés különösen hangsúlyos volt az iraki háború (2003) tervezésében, ahol a stratégiai cél a rezsim „lefejezése” volt minimális járulékos veszteséggel.<sup>15</sup>

## A technológia és a hadviselés kapcsolata clausewitzi perspektívából

Az előző alfejezetben említett doktrína azonban komoly korlátokba ütközött az iraki megszállás során. A technológiai fölény és a precíziós képességek nem tudták garantálni a politikai célok elérését, különösen az aszimmetrikus hadviselést folytató ellenállókkal szemben. Ez rávilágított arra, hogy a háború továbbra sem redukálható tisztán technikai folyamattá, és a klasszikus clausewitzi elvek – különösen a politikai

<sup>12</sup> BULEY 2008: 84–85.

<sup>13</sup> BULEY 2008: 86.

<sup>14</sup> BULEY 2008: 86.

<sup>15</sup> BULEY 2008: 116–119.

dimenzió fontossága – nem hagyhatók figyelmen kívül. A makulátlan pusztítás elmélet technológiai optimizmusa olyan önhittséget eredményezett, amely figyelmen kívül hagyta a háború inherens kiszámíthatatlanságát és káoszát – amit Clausewitz a háború „ködének és sűrűlódásának” nevezett.<sup>16</sup> A sűrűlódás (*frikcion*) Clausewitz egyik kulcsfogalma, amely a háború papíron való tervezése és tényleges végrehajtása közötti elkerülhetetlen eltérésekre utal.<sup>17</sup>

A makulátlan pusztítás elmélet szűk technikai fókuszja elhanyagolta a háború alapvetően politikai természetét és emberi dimenzióit, amelyeket az igazságos háború elmélete hagyományosan hangsúlyozott. Lényegében azáltal jelentett kihívást az igazságos háború elmélete számára, hogy azt sugallta, a technológia képes felülni a háború morális és politikai komplexitását, ami olyan feltételezésnek bizonyult, amelyet a modern hadviselés realitái nagyrészt megcáfoltak.

A technológiai fejlesztések nem csupán a katonai képességeket változtatják meg, hanem alapvető hatással vannak a háború politikai dimenziójára is. Ahogy Clausewitz hangsúlyozza, a háború a politika folytatása más eszközökkel,<sup>18</sup> azonban a mesterséges intelligencia és az autonóm rendszerek gyors, algoritmikus döntéshozatala gyakran kiszakítja a katonai műveletek irányítását és a döntéshozatalt a politikai mérlegelés és ellenőrzés hagyományos kereteiből. Az emberi felügyelet sok esetben formálissá válik a rendszerek működési sebessége és komplexitása miatt. Ez nem hagyható figyelmen kívül, amikor a politikai vezetés tényleges irányítási képességéről és felelősségéről beszélünk.

A mesterséges intelligencia nem csupán a hatékonyságot elősegítő erőmegszorozó tényezőként működik, hanem önálló taktikai és stratégiai döntéshozóként is megjelenik a modern hadviselésben.<sup>19</sup> Ez élesen jelentkezik az úgynevezett *out-of-the-loop* (emberi felügyelet nélküli) rendszerek esetében, ahol a programozó és a parancsnok közreműködése után semmiféle emberi beavatkozás nem szükséges, így ezek a rendszerek önállóan, speciális helyzetekben előre nem kiszámíthatóan viselkednek.<sup>20</sup> Ez tovább fokozza a politikai kontroll gyengülését.

## A mesterségesintelligencia-rendszerek jelentette főbb kihívások

### A feketedoboz-probléma technikai háttere

Az MI-rendszerek által felvetett kihívások közül élesen bukkan fel az elszámoltathatóság kérdése. Ki a felelős egy mesterségesintelligencia-rendszer által javasolt katonai csapásért: a parancsot kiadó tiszt, a rendszer fejlesztői vagy a politikai vezetés?

<sup>16</sup> BULEY 2008: 85.

<sup>17</sup> CLAUSEWITZ 2013 [1832].

<sup>18</sup> CLAUSEWITZ 2013 [1832]: 54.

<sup>19</sup> BODA 2024: 120.

<sup>20</sup> BODA 2022: 103.

A hagyományos parancsnoki lánc és felelősségi struktúra nehezen alkalmazható olyan helyzetekben, ahol a döntéshozatal részben átláthatatlan algoritmusokra támaszkodik. A mesterségesintelligencia-rendszerek feketedoboz-problémája alapjaiban érinti ezeket a követelményeket. A feketedoboz jellegű MI-rendszerek működési logikája különbözik a hagyományos algoritmusoktól, mivel a mély tanulási hálózatok (*deep learning networks*) több száz vagy akár ezer rejtett réteget is tartalmazhatnak. E rendszerek belső működése még a fejlesztők számára sem teljesen értelmezhető – habár látják a bemeneti adatokat és a kimeneti eredményeket, a köztes „rejtett rétegekben” zajló döntési folyamatok nem átláthatók. A legfejlettebb nyelvi modellek, mint például a ChatGPT vagy a Llama esetében a komplexitás olyan fokú, hogy még a rendszerek alkotói sem tudják pontosan megmagyarázni, hogyan jut a rendszer az adott következtetésre.<sup>21</sup> Ez a technikai korlát kérdéseket vet fel az elszámoltathatóság és a morális felelősség megállapíthatósága kapcsán, különösen a katonai alkalmazások területén, ahol a döntések potenciálisan életeket érinthetnek.

## A döntéshozatal és az irányítás átalakulása

A NATO tudományos és technológiai trendekkel kapcsolatos, a 2020–2040-es időszakra vonatkozó dokumentuma több olyan kulcsfontosságú kérdést azonosít, amelyek a mesterségesintelligencia-rendszerek katonai alkalmazásának legitimitása szempontjából meghatározók. A dokumentum kiemeli, hogy kritikus fontosságú az MI-alapú döntéstámogató rendszerek megfelelő ellenőrzése, validációja és akkreditációja (VV&A). Különösen a szövetséges műveletek során, ahol az eltérő nemzeti szabványok és adatkezelési előírások kihívást jelentenek.<sup>22</sup>

A jelentés hangsúlyozza az ember-gép együttműködés (*human-machine symbiosis*) fontosságát, ami azt jelenti, hogy az emberi döntéshozóknak és a mesterségesintelligencia-rendszereknek megbízható partnerként kell együttműködniük egy komplex hibrid rendszerben. Ez a megközelítés összhangban van azzal az elvvel, hogy a végső döntési jogkör és felelősség az emberi parancsnokoknál marad.<sup>23</sup>

Ez a követelmény ugyanakkor egyre összetettebb kihívássá válik, ahogy az MI-rendszerek fejlődése átrendezi a legitím autoritás hagyományos kereteit. A *Foreign Affairs* elemzése szerint ha a mesterséges intelligencia gyakorlatilag független politikai, diplomáciai és katonai entitásként jelenik meg, az az évszázados hatalmi egyensúlyt egy új, feltérképezetlen egyensúlytalanságra kényszerítené. Ez az átrendeződés több szinten is megkérdőjelezi a legitím autoritás hagyományos értelmezését. Egyrészt új szereplők – elsősorban technológiai vállalatok – jelennek meg, amelyek totalizáló társadalmi, gazdasági, katonai és politikai hatalom birtokosaivá válhatnak. Másrészt az MI-rendszerek fejlesztése és működése olyan zárt környezetben történik, amely kihívás elé állítja az átláthatóság követelményét.<sup>24</sup>

<sup>21</sup> KOSINSKI 2024.

<sup>22</sup> NATO Science & Technology Organization 2020: 57.

<sup>23</sup> NATO Science & Technology Organization 2020: 52.

<sup>24</sup> KISSINGER–SCHMIDT–MUNDIE 2024.

## Empirikus elemzés

Az ukrajnai háború példája különös aktualitást ad a kérdésnek. Mint azt a közelmúltban a *The New York Times* feltárta, Ukrajna gyakorlatilag „háborús techlaboratóriummá” vált, ahol az autonóm fegyverrendszerek széles körű alkalmazása zajlik.<sup>25</sup> Az amerikai katonai gondolkodásban megjelenő makulátlan pusztítás koncepció szerint a technológia lehetővé teszi a precíz, kontrollált és „tisztá” hadviselést. Ez az elképzelés azonban komoly kérdéseket vet fel a legitím autoritás elvének gyakorlati alkalmazásában.

## Ukrajnai esettanulmányok

Az ukrajnai háború egyedülálló betekintést nyújt a mesterségesintelligencia-alapú autonóm fegyverrendszerek alkalmazásának gyakorlati kihívásaiba. A *The New York Times* szerint Ukrajna gyakorlatilag „élő laboratóriummá” vált az MI-alapú hadviselésnek, ahol a technológiai újításokat valós harci körülmények között tesztelik.

Az ukrajnai konfliktusban az MI hadviselésben betöltött szerepe jelentős átalakuláson megy keresztül. A *The New York Times* részletes beszámolója alapján számos ukrán vállalat fejleszt olyan autonóm drónrendszereket és fegyvereket, amelyek minimalizálják az emberi beavatkozás szükségességét a célpontkiválasztás és a támadás során. A Vyriy vállalat által fejlesztett drónok például képesek önállóan követni és azonosítani célpontokat, míg a Saker cég olyan rendszereket alkotott, amelyek eredetileg gyümölcsök válogatására kifejlesztett MI-folyamatokat alkalmaznak célpont-azonosításra.

A fejlesztések jelenlegi szakaszában még érvényesül az úgynevezett *human in the loop* elv – a drónok például csak akkor támadhatnak, ha egy pilóta előzetesen rögzítette a célpontot. Ez az elv, amelyre gyakran hivatkoznak politikai döntéshozók és mesterségesintelligencia-etikusok, egyelőre biztosítja az emberi felügyeletet a rendszer működése felett. Ugyanakkor a technológiai fejlődés már most is lehetővé teszi az emberi beavatkozás jelentős csökkentését – példa erre a Swarmer vállalat által kifejlesztett szoftver, amely egyetlen technikus számára teszi lehetővé akár hét drón egyidejű irányítását felderítési és csapásmérési célú küldetések során. A Szerhij Kuprienko, a Swarmer alapítója és menedzsere által vezetett fejlesztés azt is lehetővé teszi, hogy a drónok akár 56 kilométeres távolságban is végrehajtsanak önálló támadásokat.

A fejlesztések mögött álló technológia nagyrészt kereskedelmi forgalomban kapható eszközökön alapul, beleértve a hobbiszámítógépeket és egyszerű alkatrészeket. Ez a megközelítés lehetővé teszi az alacsony költségű (néhány ezer dolláros) fejlesztést és a gyors adaptációt. A Saker például havonta 1000 vezérlőpanelt gyárt, és tervezi a termelés 9000 darabra növelését. Az autonóm rendszerek már bizonyították hatékonyságukat a harc téren – egy dokumentált esetben egy felderítő drón mesterséges intelligencia segítségével azonosított célpontokat, majd önálló kamikaze drónokat irányított a megsemmisítésükre 40 kilométeres távolságból.

<sup>25</sup> MOZUR–SATARIANO 2024.

A technológiai fejlődés további példája a DevDroid által fejlesztett Wolly nevű automata géppuskarendszer, amely PlayStation kontrollerral és tablettel irányítható, és képes 1000 méteres távolságból célpontokat azonosítani és követni. Habár jelenleg emberi jóváhagyás szükséges a tüzeléshez, a gyártók elismerik, hogy technikailag egyszerű lenne teljesen autonóm rendszerré alakítani.

Mihajlo Fedorov, Ukrajna digitális transzformációért felelős minisztere hangsúlyozza a maximális automatizálás szükségességét, amit az ország győzelme szempontjából alapvetőnek tart. Ez az álláspont jól tükrözi, hogy a háborús helyzetben az etikai megfontolások másodlagossá válnak a közvetlen katonai szükségletekhez képest. Stuart Russell, a Berkeley Egyetem mesterségesintelligencia-kutatója szerint ez a fejlődés egyértelművé tette az autonóm fegyverek előnyeit, ugyanakkor arra figyelmeztet, hogy ezek a technológiák könnyen hozzáférhető tömegpusztító fegyverekké válhatnak a globális fegyverpiacon.<sup>26</sup>

## A Palantir példája

A Palantir vállalat ukrainai tevékenysége példaértékű esettanulmányt kínál a mesterséges intelligencia katonai döntéshozatalban betöltött szerepének elemzéséhez, különös tekintettel a parancsnoki jogkörök és felelősségek átalakulására. A *Time* magazin részletes beszámolója alapján a vállalat olyan összetett MI-rendszert telepített Ukrajnában, amely többféle adatforrást – műholdképeket, drónfelvételeket, földi hírszerzési jelentéseket, felhőzetten átható radaradatokat és hőérzékelő képeket – integrál és elemez. A rendszer ezeket az információkat mesterséges intelligencia segítségével dolgozza fel, majd konkrét katonai cselekvési opciókat javasol a parancsnokoknak.

Alex Karp, a Palantir vállalat vezérigazgatójának állítása szerint rendszereik felelősek az ukrainai célpontok többségének kiválasztásáért, ami jelzi a technológia meghatározó szerepét. A rendszer példátlan mértékű integrációja – amely kiterjed a Védelmi Minisztériumtól az Oktatási Minisztériumig számos kormányzati szervre – új helyzetet teremt a katonai döntéshozatalban. A zászlóaljokhoz telepített szoftvermérnökök jelenléte pedig azt mutatja, hogy a kereskedelmi technológia közvetlenül beépült a harctéri műveletekbe.

Ez a gyakorlat több szempontból is kihívást jelent a hagyományos parancsnoki struktúrák számára. Egyrészt a rendszer működése nem teljesen átlátható a katonai vezetők számára, mivel a döntési algoritmusok részben feketedobozként működnek. Másrészt a szoftver – a vállalat állítása szerint – minden csapással tanul és fejlődik, ami felveti a kérdést, ki felügyeli és irányítja ezt a fejlődési folyamatot. A helyzetet tovább bonyolítja, hogy egy kereskedelmi vállalat tulajdonában lévő technológia vált a katonai műveletek szerves részévé, létrehozva azt, amit Karp digitalizált célmegsemmisítési láncnak (*digital kill chain*) nevez.

Ezt az átalakulást Mark Milley tábornok, volt amerikai vezérkari főnök a hadviselés jellegének a történelemben a legjelentősebb alapvető változásaként említette.<sup>27</sup>

<sup>26</sup> MOZUR–SATARIANO 2024.

<sup>27</sup> BERGENGRUEN 2024.

Mindez számos stratégiai kérdést is felvet: idetartozik a kereskedelmi technológiától való függőség problémája, az üzleti és katonai célok közötti lehetséges érdekkellentétek, valamint az MI-rendszerek által befolyásolt célpontkiválasztás elszámoltathatóságának kérdése. Az ukrajnai példa így nemcsak a jelenlegi konfliktus szempontjából jelentős, hanem a jövőbeli katonai-kereskedelmi együttműködések és a mesterséges intelligencia által vezérelt hadviselés tesztjeként is szolgál.

## Elemzés

Az empirikus példák alapján most visszatérünk a tanulmány elején feltett kérdésekhez, és megvizsgáljuk, hogyan értelmezhető az elszámoltathatóság a feketedoboz-probléma tükrében, milyen szerepet tölt be a *human in the loop* elv a legitím autoritás fenntartásában, és milyen lehetőségek kínálóznak a döntéshozatal átláthatóságának biztosítására az MI-alapú rendszerek esetén.

Az empirikus példák értelmében a legitím autoritás elve három kulcsfontosságú területen ütközik kihívásokba a mesterségesintelligencia-alapú rendszerek alkalmazása során:

- Először is, a döntéshozatali folyamat átláthatósága jelentősen csökken. A feketedoboz-probléma nem pusztán technikai kérdés, hanem alapvetően érinti a katonai vezetés legitimitását. Amikor egy mesterségesintelligencia-alapú rendszer javasol célpontokat vagy taktikai opciókat, a parancsnokok gyakran nem érthetik teljes mértékben a javaslat mögötti logikát. Ez ellentmond az igazságos háború elmélete azon alapelveinek, hogy a hadviselésben hozott döntéseknek racionálisan indokolhatónak és morálisan védhetőnek kell lenniük.<sup>28</sup>
- Másodsor, az elszámoltathatóság hagyományos mechanizmusai felbomlanak. A parancsnoki lánc, amely tradicionálisan biztosította a felelősség egyértelmű megállapíthatóságát, most összetett ember-gép interakciók hálózatává válik. Mint azt az ukrajnai példák mutatják, a döntéshozatalban részt vevő szereplők köre kibővül kereskedelmi vállalatokkal és azok MI-alapú rendszereivel, ami tovább bonyolítja a felelősségi viszonyokat.
- Harmadsor, a *human in the loop* elv, habár fontos biztosítéknak tűnik, a gyakorlatban gyakran nem képes teljes mértékben garantálni az érdemi emberi kontrollt. A modern hadviselés tempója és információs komplexitása olyan helyzeteket teremt, ahol az emberi felügyelet könnyen formálissá válhat. Ez különösen problematikus az igazságos háború elmélete szempontjából, amely feltételezi az erkölcsi mérlegelés lehetőségét minden jelentős katonai döntésnél.

A legitím autoritás elvének újraértelmezését jelentősen befolyásolhatják a legújabb technológiai fejlemények. Az Anthropic mesterségesintelligencia-laboratórium 2024 májusában bejelentett áttörése a feketedoboz-probléma potenciális megoldását ígéri. A kutatók olyan technikát fejlesztettek ki, amely lehetővé teszi az MI-modellek „agyának” szkennelését, azonosítva azokat a neuroncsoportokat, amelyek különböző

<sup>28</sup> WALZER 2015.

koncepciókhoz kapcsolódnak.<sup>29</sup> Ez a fejlesztés új perspektívát nyithat a legitím autoritás elv MI-rendszerek alkalmazásában való érvényesülésének vizsgálatában.

Ugyanakkor fontos megjegyezni, hogy még ez az áttörés sem jelenti a teljes átláthatóság elérését. Mint azt az Anthropic kutatói jelzik, a rendszer összes jellemzőjének azonosítása még mindig prohibitív költségekkel járna. Ráadásul a biztonsági szempontból releváns jellemzők manipulálásának hosszú távú hatásai további vizsgálatot igényelnek.

Ez a technológiai fejlődés tehát nem oldja fel teljesen a legitím autoritás elvével kapcsolatos alapvető dilemmákat, de új eszközöket kínálhat az elszámoltathatóság és az átláthatóság javítására. A neuroncsoportok azonosítása és manipulálása például lehetővé teheti a biztonsági kockázatok jobb kezelését, de továbbra is nyitott kérdés marad, hogy ki és milyen felhatalmazással végezhet ilyen beavatkozásokat.

A katonai döntéshozatal legitimitásának és elszámoltathatóságának kérdése szorosan összefügg a mesterségesintelligencia-rendszerek általános etikai problémáival. Ahogy Floridi kifejti, az MI-rendszerek alapvető működési elve, hogy elválasztják a feladat sikeres végrehajtásának képességét az intelligencia szükségességétől.<sup>30</sup> Ez a katonai alkalmazások szempontjából kritikus kérdéseket vet fel, hiszen egy olyan döntéshozó rendszer legitimitása lehet kérdéses, amely nem képes döntései következményeinek teljes körű megértésére.

Az elszámoltathatóság technikai feltételei ugyanakkor folyamatosan fejlődnek. Egyre több megfelelő eszköz áll rendelkezésre a gépi tanulási rendszerek kimeneteinek megfigyelésére és megértésére.<sup>31</sup> Ez különösen fontos a katonai döntések esetében, ahol alapvető követelmény az ellenőrizhetőség és az átláthatóság.

A felelősség kérdésében Floridi határozott álláspontot képvisel: a mesterséges intelligenciát normál technológiaként kell kezelni. Nem csoda és nem is csapás, csupán egy a sok megoldás közül, amit az emberi találékonyság létrehozott. Ezért is marad az etikai vita örökre tisztán emberi kérdés.<sup>32</sup> Ez azt sugallja, hogy noha az MI-rendszerek alkalmazhatók katonai döntéstámogatásra, a végső felelősség mindig emberi kézben kell hogy maradjon.

A Floridi által bevezetett különbségtétel a „komplex” és „nehéz” problémák között<sup>33</sup> további támpontot nyújt annak megítéléséhez, hogy mely döntések delegálhatók MI-rendszerekhez. Ez a keretrendszer segíthet meghatározni azokat a területeket, ahol a mesterséges intelligencia támogathatja a katonai döntéshozatalt, miközben fenntartja az emberi kontrollt és felelősséget.

<sup>29</sup> PERRIGO 2024.

<sup>30</sup> FLORIDI 2023: 39.

<sup>31</sup> FLORIDI 2023: 51.

<sup>32</sup> FLORIDI 2023: 52.

<sup>33</sup> FLORIDI 2023: 39–40.

## Az elszámoltathatóság új dimenziói: *ex ante* kontroll és szabályozási keretek

### Az emberi kontroll újraértelmezése: a békeidős előkészítés modellje

A legitím autoritás kérdése az MI-alapú fegyverrendszerek kontextusában különösen aktuálissá vált, amit jól szemléltet Sebastian Elbaumnak, a Virginia Egyetem professzorának és Jonathan Panternek, a Council on Foreign Relations szakértőjének elemzése a *Foreign Affairs* hasábjain. A szerzők amellet érvelnek, hogy az emberi kontroll fenntartásának hagyományos megközelítése egyre kevésbé tartható a modern hadviselésben, ami alapvető kihívások elé állítja az igazságos háború elméletének alkalmazását.

Az elsődleges probléma az MI-rendszerek komplexitása és működési sebessége: a mesterségesintelligencia-modellek, amelyek a modern autonóm fegyvereket irányítják, gyakran túl kifinomultak ahhoz, hogy még a legjobban képzett operátorok is felügyeljék őket. Ez közvetlen kihívást jelent a legitím autoritás gyakorlása szempontjából, hiszen megkérdőjelezi az érdemi emberi döntéshozatal lehetőségét. A szerzők rámutatnak, hogy harci körülmények között – amelyeket súlyos stressz, korlátozott idő és megszakított vagy nem létező kommunikáció jellemez az egyének, egységek és magasabb hatóságok között – gyakorlatilag lehetetlen az értelmes emberi beavatkozás. Ez a megfigyelés összhangban áll Walzer azon megállapításával, hogy a legitím autoritás gyakorlásának módja szorosan összefügg a hadviselés aktuális körülményeivel. A megoldást a szerzők szerint az jelentheti, ha az etikai kereteket és a döntési paramétereket békeidőben határozzák meg: a kormányoknak először azt kell biztosítaniuk, hogy a kulcsfontosságú emberi döntések, mint például az arányosság és a megkülönböztetés, ne háborús időben szülessenek, amikor nem lesz idő gondolkodni, hanem békeidőben, amikor szakértői inputra, demokratikus vitára és magas szintű felügyeletre van lehetőség. Ez a megközelítés új perspektívát nyit a legitím autoritás gyakorlásának időbeli dimenziójára vonatkozóan.

A szerzők hangsúlyozzák, hogy az autonóm fegyverrendszerek fejlesztése és telepítése gyakorlatilag elkerülhetetlen a nagyhatalmak közötti verseny kontextusában. Ez a megállapítás fontos az igazságos háború elmélete szempontjából, mivel felveti a kérdést, hogy a technológiai fejlődés és a geopolitikai realitások hogyan befolyásolják az igazságos háború kritériumainak értelmezését és alkalmazását.<sup>34</sup>

Elbaum és Panter elemzése tehát azt sugallja, hogy a legitím autoritás gyakorlásának módja jelentős átalakuláson megy keresztül az MI-alapú fegyverrendszerek korában. A hagyományos *human in the loop* megközelítés helyett olyan új struktúrára és folyamatokra van szükség, amelyek békeidőben biztosítják az etikus kereteket és a megfelelő felügyeletet, miközben lehetővé teszik a rendszerek hatékony működését harci körülmények között.

<sup>34</sup> ELBAUM–PANTER 2024.

## Nemzetközi humanitárius jogi keretek

A nemzetközi humanitárius jog keretrendszere megfontolást követel a mesterségesintelligencia-rendszerek alkalmazása során. Mint azt Elliot Winter nemzetközi jogi professzor kifejti, a megkülönböztetés elve (*distinction principle*) és az arányosság követelménye például olyan kontextusérzékeny mérlegelést igényel, amit az MI-rendszerek még nem képesek megfelelően végrehajtani. A gépi megfigyelés és felismerés terén elért sikerek ellenére a gépi döntéshozatal még mindig elmarad az emberi képességektől. A döntéshozatal kritikus fontosságú a nemzetközi humanitárius jog szerinti megkülönböztetés szempontjából, mivel az emberek fizikai megjelenésük változása nélkül is válhatnak a támadható kombattáns és a nem támadható civil státusz között. Ugyanakkor a technológiai fejlődés – különösen a magas szintű MI várható megjelenése – középtávon lehetővé teheti e követelmények teljesítését, hiszen valószínűsíthető, hogy 2040 és 2062 között már olyan fejlett mesterséges intelligenciával fogunk rendelkezni, amely képes lesz kezelni a megkülönböztetési értékelések teljes spektrumát.<sup>35</sup>

## Az Európai Unió mesterséges intelligenciáról szóló rendelete

Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet) célja az MI-technológiák biztonságos és az alapvető jogokat tiszteletben tartó fejlesztésének és alkalmazásának biztosítása az EU-ban, mindezt oly módon kivitelezve, hogy a szabályozás ne hátráltassa az innovációt.<sup>36</sup>

A rendelet a mesterségesintelligencia-rendszerek első átfogó szabályozási kerete, amely különös figyelmet fordít a magas kockázatú MI-alkalmazásokra. Habár a katonai felhasználás explicit módon kivételt képez a szabályozás alól, az alapelvei – különösen az átláthatóság és az elszámoltathatóság követelményei – relevánsak a legitím autoritás újraértelmezése szempontjából.

A rendelet által előírt átláthatósági kötelezettségek például párhuzamba állíthatók az igazságos háború elméletének elszámoltathatósági követelményeivel. A szabályozás megköveteli az MI-rendszerek döntéshozatali folyamatainak dokumentálását és az emberi felügyelet biztosítását, ami összhangban van az emberi felügyelet (*human in the loop*) elvvel a katonai alkalmazásokban. A szabályozás által bevezetett kockázatalapú megközelítés szintén tanulságos: minél nagyobb a potenciális kár, annál szigorúbbak a követelmények. Ez az elv adaptálható lehet a katonai mesterségesintelligencia-rendszerek esetében is, ahol a döntések következményei különösen súlyosak lehetnek.

<sup>35</sup> WINTER 2022: 14–15.

<sup>36</sup> Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete.

A rendelet kockázatalapú megközelítést alkalmaz: tiltja az elfogadhatatlan kockázatú rendszereket,<sup>37</sup> szigorúan szabályozza a magas kockázatú alkalmazásokat,<sup>38</sup> míg az alacsony kockázatú rendszerekre enyhébb átláthatósági követelményeket ír elő. A rendelet hatálya kiterjed minden MI-szolgáltatóra és felhasználóra, aki az EU piacán terméket vagy szolgáltatást kínál, függetlenül attól, hogy az EU-ban vagy harmadik országban található.

A rendelet ugyan elsődlegesen nem katonai alkalmazásokra fókuszál, de átláthatósági és elszámoltathatósági alapelvei relevánsak a katonai felhasználás szempontjából. A rendelet szerint minden magas kockázatú mesterségesintelligencia-rendszernek meg kell felelnie az átláthatósági és dokumentációs követelményeknek. A rendszerek működését nyomon követhetővé kell tenni automatikus eseménynaplózással, és biztosítani kell az emberi felügyeletet.

A magas kockázatú rendszerek esetében a szolgáltatóknak részletes technikai dokumentációt kell készíteniük,<sup>39</sup> amely bemutatja a rendszer működését és korlátait. A szolgáltatóknak kockázatkezelési rendszert kell működtetniük,<sup>40</sup> és biztosítaniuk kell a megfelelő szintű pontosságot, robusztusságot és kiberbiztonsági védelmet.<sup>41</sup>

Noha a katonai alkalmazások nem tartoznak a rendelet hatálya alá (a preambulumbekezdés 24 és 2. cikk 3. bekezdés alapján), az alapelvek – különösen az átláthatóság és az elszámoltathatóság követelménye – hasznos iránymutatást adhatnak a katonai mesterségesintelligencia-rendszerek szabályozásához is. A törvény által előírt technikai és szervezeti intézkedések, mint például az automatikus naplózás, az emberi felügyelet és a részletes dokumentáció, adaptálhatók a katonai kontextusra is.

## Összegző gondolatok

Az elemzés megerősítette kezdeti feltételezésünket: a mesterségesintelligencia-alapú autonóm fegyverrendszerek alkalmazása valóban gyengíti a legitím autoritás elvének hagyományos értelmezését. Ugyanakkor nem elegendő pusztán ennek megállapítása, hanem szükséges az elv újraértelmezése a megváltozott technológiai környezetben. A tanulmány elején feltett kérdésekre a következő válaszokat adhatjuk.

*Az elszámoltathatóság értelmezése a feketedoboz-probléma tükrében.* A mesterségesintelligencia-rendszerek átláthatatlan belső működése jelentősen nehezíti az elszámoltathatóság érvényesítését. Mint azt az ukrajnai példák is mutatják, a katonai döntéshozók gyakran nem érthetik teljes mértékben az algoritmusok javaslatainak hátterét, ami aláássa a felelősség megállapíthatóságát. Az Anthropic MI-laboratórium legújabb fejlesztései ugyanakkor reményt adnak arra, hogy a feketedoboz részleges „felnyitása” révén javítható legyen az átláthatóság és így az elszámoltathatóság is.

*Az emberi felügyelet elvének szerepe a legitím autoritás fenntartásában.* A tanulmány megállapította, hogy noha az elv formálisan biztosítja az emberi kontrollt, a modern

<sup>37</sup> Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete, 5. cikk.

<sup>38</sup> Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete, 6–17. cikk.

<sup>39</sup> Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete, 11. cikk.

<sup>40</sup> Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete, 9. cikk.

<sup>41</sup> Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete, 15. cikk.

hadviselés sebessége és komplexitása miatt ez gyakran kiüresedő mechanizmussá válik. Az ukrainai esettanulmányok rávilágítottak, hogy ezen elv gyakran inkább formális követelménynek tűnik, mintsem érdemi ellenőrzésnek. A legitím autoritás fenntartásához szükséges átgondolni e mechanizmus hatékonyabb megvalósítási módjait, különös tekintettel Elbaum és Panter javaslatára, amely szerint a kulcsfontosságú etikai döntéseket békeidőben kellene meghozni.

*A döntéshozatal átláthatóságának biztosítási lehetőségei a mesterségesintelligencia-alapú rendszerek esetén.* A tanulmány megállapította, hogy az átláthatóság technikai szintű biztosítása mellett (mint az EU mesterséges intelligenciáról szóló rendeletének követelményei) lényeges a döntéshozatal előkészítésének és utólagos értékelésének folyamata. Az *ex ante* kontroll és a megfelelő szabályozási keretek kialakítása jelentheti azt az utat, amelyen haladva a mesterségesintelligencia-rendszerek alkalmazása összeegyeztethetővé válik a legitím autoritás elvének lényegi követelményeivel.

A legitím autoritás elvének adaptációja során három kulcsfontosságú szempont tűnik megfontolandónak. Először is, az átláthatóság követelményének újragondolása: míg a hagyományos értelmezés teljes transzparenciát feltételezett, az MI-rendszerek esetében ez technikailag nem megvalósítható; ehelyett olyan új standardokra van szükség, amelyek biztosítják a döntéshozatal megfelelő szintű ellenőrizhetőségét a feketedoboz-probléma ellenére is. Másodsorban, az elszámoltathatóság új modelljének kidolgozása: a tradicionális parancsnoki lánc helyett egy összetettebb felelősségi rendszer kialakítása, amely figyelembe veszi az ember-gép interakciók komplexitását; ez magában foglalhatja a fejlesztők, a katonai vezetők és a politikai döntéshozók közötti felelősség megosztásának új formáit. Végül az emberi felügyelet elvének megerősítése: nem elégséges ennek formális biztosítása; olyan mechanizmusokra van szükség, amelyek garantálják az érdemi emberi kontrollt a kritikus döntések felett.

További elemzést igényel annak vizsgálata, hogy ezek az adaptációk hogyan valósíthatók meg a gyakorlatban, különös tekintettel a gyorsan fejlődő technológiai környezetre. Az igazságos háború elmélet legitím autoritás elvének újraértelmezése nem csupán elméleti kérdés, hanem sürgető gyakorlati szükségszerűség a modern hadviselésben.

Végezetül fontos megjegyezni néhány kritikus szempontot az igazságos háború elmélete és a mesterséges intelligencia kapcsolatának további vizsgálatához:

- Többféle elmélet létezik az igazságos háborúra, amelyek közül Walzer megközelítése csak egy, de meghatározó értelmezés. A jelen tanulmány elsősorban az ő elméletére támaszkodott, de a téma teljesebb megértéséhez más megközelítések vizsgálata is szükséges lehet.
- Míg vannak konkrét esetek a mesterséges intelligencia harctéri alkalmazására, a katonai döntéshozatalban való alkalmazás még kevésbé dokumentált terület. Ugyanakkor a technológiai fejlődés üteme alapján várható, hogy hamarosan ezen a téren is jelentős változások következnek be.
- Az MI-rendszerek katonai alkalmazásának szabályozása továbbra is elsősorban nemzetállami hatáskörben marad, amit az is megerősít, hogy az EU mesterséges intelligenciáról szóló rendelete a kizárólag katonai célú alkalmazásokat kiveszi a hatálya alól. Ugyanakkor az átláthatóság és az elszámoltathatóság általános elvei iránymutatást adhatnak a katonai fejlesztések során is.

A legitím autoritás elvének újraértelmezése a mesterséges intelligencia korában tehát nemcsak technológiai és jogi, hanem alapvető etikai kérdéseket is felvet, amelyek további elemzést és nemzetközi párbeszédet igényelnek.

## Felhasznált irodalom

- BERGENGRUEN, Vera (2024): How Tech Giants Turned Ukraine Into an AI War Lab. *Time*, 2024. február 8. Online: <https://time.com/6691662/ai-ukraine-war-palantir/>
- BODA Mihály (2022): A hibrid háború etikája: az igazságos hibrid háború elmélete. In M. SZABÓ Miklós (szerk.): *A hadtudomány aktuális kérdései I.* Budapest: Ludovika Egyetemi Kiadó, 95–108.
- BODA Mihály (2023): Az igazságos háború elmélete és a teremtésvédelem: Az igazságos háború elméletének története a középkortól a 20. századig a teremtésvédelem szempontjából. *Hadtudományi Szemle*, 16(4), 123–137. Online: <https://doi.org/10.32563/hsz.2023.4.6>
- BODA Mihály (2024): A kockázatkerülő háború és a bátorság a 20–21. század fordulóján. *Honvédségi Szemle*, 152(3), 113–125. Online: <https://doi.org/10.35926/HSZ.2024.3.9>
- BULEY, Benjamin (2008): *The New American Way of War: Military Culture and the Political Utility of Force*. London: Routledge Taylor & Francis Group. Online: <https://doi.org/10.4324/9780203934166>
- CLAUSEWITZ, Carl von (2013) [1832]: *A háborúról*. Budapest: Zrínyi Kiadó.
- ELBAUM, Sebastian – PANTER, Jonathan (2024): AI Weapons and the Dangerous Illusion of Human Control. *Foreign Affairs*, 2024. december 6. Online: <https://www.foreignaffairs.com/united-states/ai-weapons-and-dangerous-illusion-human-control>
- FLORIDI, Luciano (2023): *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford: Oxford University Press. Online: <https://doi.org/10.1093/oso/9780198883098.001.0001>
- JOHNSON, James (2022): The AI Commander Problem: Ethical, Political, and Psychological Dilemmas of Human-Machine Interactions in AI-enabled Warfare. *Journal of Military Ethics*, 21(3–4), 246–271. Online: <https://doi.org/10.1080/15027570.2023.2175887>
- KISSINGER, Henry A. – SCHMIDT, Eric – MUNDIE, Craig (2024): War and Peace in the Age of Artificial Intelligence. *Foreign Affairs*, 2024. november 18. Online: <https://www.foreignaffairs.com/united-states/war-and-peace-age-artificial-intelligence>
- KOSINSKI, Matthew (2024): What Is Black Box AI? *IBM*, 2024. október 29. Online: <https://www.ibm.com/think/topics/black-box-ai>
- MOZUR, Paul – SATARIANO, Adam (2024): A.I. Begins Ushering In an Age of Killer Robots. *The New York Times*, 2024. július 2. Online: <https://www.nytimes.com/2024/07/02/technology/ukraine-war-ai-weapons.html>
- NATO Science & Technology Organization (2020): *Science & Technology Trends: 2020–2040*. Online: [https://www.nato.int/content/dam/nato/legacy-wcm/media\\_pdf/2020/4/pdf/190422-ST\\_Tech\\_Trends\\_Report\\_2020-2040.pdf](https://www.nato.int/content/dam/nato/legacy-wcm/media_pdf/2020/4/pdf/190422-ST_Tech_Trends_Report_2020-2040.pdf)

- NOLAN, Cynthia (2015): Michael Walzer. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. 5th ed. New York: Basic Books, 2015. Book Review. *Saber and Scroll Journal*, 4(3), 125–127.
- PERRIGO, Billy (2024): No One Truly Knows How AI Systems Work. A New Discovery Could Change That. *Time*, 2024. május 21. Online: <https://time.com/6980210/anthropic-interpretability-ai-safety-research/>
- WALZER, Michael (2015): *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. New York: Basic Books.
- WINTER, Elliot (2022): The Compatibility of Autonomous Weapons with the Principles of International Humanitarian Law. *Journal of Conflict and Security Law*, 27(1), 1–20. Online: <https://doi.org/10.1093/jcsl/krac001>

## Jogi forrás

Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet). Online: <https://eur-lex.europa.eu/legal-content/HU/ALL/?uri=CELEX%3A32024R1689>