

УДК 341.3

DOI <https://doi.org/10.24144/2307-3322.2026.93.5.46>

ОБМЕЖЕННЯ ТА ВІДПОВІДАЛЬНІСТЬ ЗА ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У ЗБРОЙНИХ КОНФЛІКТАХ: НАЦІОНАЛЬНИЙ І МІЖНАРОДНИЙ ВИМІРИ

Каряка Р.О.,

*аспірант кафедри міжнародного права та порівняльного правознавства
Чернівецького національного університету*

ім. Ю. Федьковича

ORCID: 0009-0008-0673-4423

Каряка Р.О. Обмеження та відповідальність за використання штучного інтелекту у збройних конфліктах: національний і міжнародний виміри.

У статті досліджуються існуючі міжнародні та національні підходи до обмеження штучного інтелекту, що використовується у військовій промисловості. Ставиться під питання, чи забезпечують Женевські конвенції достатню кількість запобіжних заходів для суб'єктів міжнародного гуманітарного права, які мають наміри застосовувати автономні системи під час збройних конфліктів. Розглянуто історію становлення поняття «штучний інтелект» на прикладі національних теорій науковців.

Проведено паралелі правосуб'єктності особи та комп'ютера, наділеного розумовими здібностями, мисленням та поведінкою людини, задля з'ясування можливості встановлення відповідальності для штучного інтелекту. Встановлено, що «штучність» інтелекту у військових системах не є автономією в розумінні автономії волі людини. А правосуб'єктність штучного інтелекту виключає можливість притягнення комп'ютера до відповідальності у зв'язку з відсутністю правового бачення його як повноцінного суб'єкта правовідносин.

Детально розглянуто питання відповідальності за використання штучного інтелекту у збройних конфліктах в межах міжнародних суб'єктів – держав та міжнародних організацій, а також національних – командирів, виробників, розробників та операторів.

За результатами дослідження зроблено висновок, що обмеження використання штучного інтелекту має зводитися до прямої заборони алгоритмічного мислення у новітній зброї без прямого контролю людини. Щоб можна було говорити про міжнародно-правову відповідальність в порушення норм права війни комбатант має лишатися єдиним суб'єктом притягнення. А штучний інтелект у системах озброєнь мусить лишатися неавтономним джерелом інформації для проведення аналізу поля бою оператором або командиром. Як-от використання штучного розуму у системах протиповітряної оборони, у супутниках чи дронах розвідників. Будь-яка смертельна зброя за участі штучного інтелекту обмежується моделлю «human-in-the-loop», а прийняття державами Женевських конвенцій та додаткових протоколів до них встановили запобіжні заходи щодо використання автономної зброї під час збройних конфліктів.

Ключові слова: штучний інтелект, міжнародне гуманітарне право, автономні системи озброєння, військові злочини, юридична відповідальність, право збройних конфліктів.

Kariaka R.O. Restrictions and liability for the use of artificial intelligence in armed conflicts: national and international dimensions.

The article examines existing international and national approaches to restricting artificial intelligence used in the military industry. It questions whether the Geneva Conventions provide a sufficient number of safeguards for subjects of international humanitarian law who intend to use artificial intelligence during armed conflicts. The historical background of the concept of «artificial intelligence» is considered on the example of national theories of scientists.

Parallels are drawn between the legal personality of an individual and a computer endowed with mental abilities, thinking, and human behavior in order to clarify the possibility of establishing liability

for artificial intelligence. It has been established that the «artificiality» of intelligence in military systems is not autonomy in the sense of human autonomy of will. And the legal personality of artificial intelligence excludes the possibility of holding a computer liable due to the lack of a legal vision of artificial intelligence as a full-fledged subject of legal relations.

The issue of liability for the use of artificial intelligence in armed conflicts within the framework of international subjects – states and international organizations, as well as national ones – commanders, manufacturers, developers, and operators, is considered in detail.

Based on the results of the research, it is concluded that restrictions on the use of artificial intelligence should be reduced to a direct ban on algorithmic thinking in modern weapons without direct human control. In order to speak about international legal liability for violation of the laws of war, the combatant must remain the only subject of prosecution. And artificial intelligence in weapon systems must remain a non-autonomous source of information for battlefield analysis by an operator or commander. Such as the use of artificial intelligence in air defense systems, satellites, or reconnaissance drones. Any lethal weapon involving artificial intelligence is limited by the «human-in-the-loop» model, and the adoption by states of the Geneva Conventions and their Additional Protocols established safeguards regarding the use of autonomous weapons during armed conflicts.

Key words: artificial intelligence, international humanitarian law, autonomous weapon systems, war crimes, legal liability, law of armed conflict.

Постановка проблеми. У зв'язку із поширенням застосування штучного інтелекту серед армій світу в безпілотних літальних апаратах, системах протиповітряної оборони, системах спостереження, ракетах, великого значення набуває питання повної заборони чи обмеження автономних систем, що несуть загрозу цивільному населенню. І якщо сам факт застосування може обмежуватися людським фактором, то межі такого використання, а також розрізнення відповідальності осіб за впливом на театр воєнних дій виокремлюються в окремий блок регулювання міжнародного гуманітарного права. Тому штучний інтелект у воєнній промисловості матиме дискусійний характер до моменту однакового розуміння сторонами збройних конфліктів меж застосування штучного розуму, а також доцільності використання такого алгоритму дій з огляду на непередбачуваність наслідків.

До того ж, зростання геополітичної напруги та щорічне збільшення конфліктів змінює сучасну парадигму правил ведення війни. Вже зараз приватні компанії створюють спеціалізовані інструменти для збройних сил держав задля обробки великої кількості інформації з метою пришвидшення ухвалення військових рішень. В свою чергу механізми регулювання такої діяльності не розроблені у правовому полі. Перегони озброєнь в найближчому майбутньому вимагатимуть встановлення чітких обмежень, регуляцій та окремого інституту відповідальності за використання штучного інтелекту у збройному конфлікті.

Мета дослідження: комплексний аналіз міжнародно-правових та національних механізмів регулювання використання штучного інтелекту під час збройних конфліктів задля визначення найбільш дієвих моделей його обмеження, а також виокремлення фундаментальних правових засад для формування концепції інституту відповідальності за використання таких систем у міжнародному гуманітарному праві.

Стан опрацювання проблематики. Питання правового регулювання штучного інтелекту у сфері національної безпеки та забезпечення основоположних прав людини під час збройного конфлікту розглядаються вітчизняними науковцями з різних сторін. Окремі проблеми досліджуваного питання були предметом розгляду Гачкевич А.О., Татаренко Г.В., Волобоєва А.О., Шут В.О., Ганенко І.С., та більш ранніх іноземних науковців Тюрінга А. та МакКарті Д. Проте комплексне вивчення існуючих регулятивних інструментів у сучасному міжнародному гуманітарному праві та національному законодавстві держав щодо відповідальності за використання штучного інтелекту у збройних конфліктах та обмеження таких систем залишається недостатнім.

Виклад основного матеріалу. Поняття штучного інтелекту за своїм розумінням почало поширюватися у 1950-х роках, зокрема серед англійських дослідників, що формували перші теорії про можливість копіювання електронними машинами розуму та поведінки людини. Одними з перших в цьому розвитку відзначилися Алан Тюрінг – англійський математик, логік і криптограф, та Джон МакКарті – американський інформатик та дослідник мислення. Одноіменний «Тест Тюрінга» закріпив у теорії комп'ютерного мислення розуміння того, що залучення машини у життя

людини завжди буде обмежуватися набором закладених у неї параметрів та вихідних даних. Суть полягала в тому, що сам тест був грою у визначення фальшивості тверджень щодо статі учасника експерименту або їх істинності. Визначалося 3 учасники, серед яких був спостерігач (або суддя), який фактично вгадує хто є чоловіком (учасник 2), а хто жінкою (учасник 3). Згодом в гру додавався комп'ютер, який надавав відповіді на запитання спостерігача, враховуючи закладену в нього інформацію, для того, щоб видати себе за людину [1, с. 65–66]. Тобто таке мислення по суті є штучним з огляду на відсутність у комп'ютера когнітивних функцій наряду з мораллю та духовністю.

В той же час, у 1956 році Джон МакКарті вперше вводить в інформатичну науку поняття «штучного інтелекту» і зазначає, що таке мислення не слід розуміти як пряме копіювання думок і поведінки людини, інтелект здатен використовувати власні методи комунікації та світосприйняття [2, с. 77]. Це розуміння перекликається з «Тестом Тюрінга», враховуючи, що під час гри комп'ютер здатний вводити в оману наглядача, допускаючи помилки при вирішенні математичних рівнянь або інших задач із числами, де допускається «людський фактор».

Оминаючи моральну та релігійну складову кожного збройного конфлікту, військово-використання штучного інтелекту відсилає до первинних джерел міжнародного гуманітарного права. Зокрема, Додатковим протоколом I до Женевських конвенцій (далі – Додатковий протокол I) визначається два основних обмеження щодо використання штучного інтелекту у новітній зброї, а саме: принцип розрізнення цивільного населення та комбатантів і обмеження нової зброї без прямої правової заборони [3, с. 229]. Згідно зі статтею 48 вказаного протоколу сторони конфлікту зобов'язуються розрізняти комбатантів та цивільне населення, в тому числі вживаючи постійних запобіжних заходів (стаття 57) із заборонаю невибіркових нападів (стаття 51) [4]. Тобто встановлюється пряма залежність від когнітивних функцій людини щодо розрізнення і прийняття рішення до можливості нести відповідальність за вказані дії.

Як зазначалося вище, способи ведення війни обмежуються Додатковим протоколом I. В свою чергу даний нормативно-правовий акт обмежує сторони конфлікту у виборі методів та засобів, що застосовуються у збройному конфлікті (стаття 35), а при розробці і використанні новітньої зброї держави зобов'язуються проводити юридичну експертизу, щоб визначити, чи підпадає вона під заборони міжнародного гуманітарного права (стаття 36) [5, с. 139]. Відтак, використання штучного інтелекту де-юре обмежується Правом Женеви без наявності спеціального міжнародного нормативно-правового акту.

Варто також зазначити, що міжнародне гуманітарне право забезпечується дотриманням принципу пропорційності, який співвідноситься з принципом розрізнення. Якщо автоматизована зброя обирає за ціль лише військових – принципи дотримано, однак якщо штучний інтелект не зміг розрізнити цивільний та військовий об'єкт і завдав неправомірної шкоди, принцип порушено. До того ж, автономна система не буде здатна розпізнати серед комбатантів осіб, які більше не беруть участь у воєнних діях: полонені, поранені та хворі, особи, які явно виявляють намір здатися (*hors de combat*) [4].

Додатковою складністю для штучного інтелекту виступає закріплений в міжнародному гуманітарному праві принцип військової необхідності. Він перш за все забезпечується шляхом активного втручання сторін конфлікту в перебіг бойових дій, коли з'являється потреба в ослабленні супротиву противника чи руйнування планів. Зразковими приписами Додаткового протоколу I виступає стаття 52, де визначаються умови, за яких цивільний об'єкт може стати законною військовою ціллю. Або за статтею 51 сторони повинні враховувати під час бойових дій необхідність завдання атаки, що може спричинити додаткові непропорційні втрати серед цивільного населення. В той же час, стаття 57 встановлює обов'язок сторін оберігати цивільне населення та об'єкти [4]. З чого слідує, що вихідні положення принципу військової необхідності будуються на персональній відповідальності комбатанта, що приймає таке рішення. До того ж, закладення в автономну систему детальних інструкцій не може гарантувати пропорційність кожного рішення, оскільки людина, приймаючи рішення, покладається не тільки на об'єктивні, але також і на суб'єктивні фактори.

Враховуючи, що відповідальність корелюється з правосуб'єктністю, яка визначає здатність людини виступати учасником правовідносин, штучний інтелект, як механізм із шаблонним людським мисленням, поведінкою та методами комунікації, міг би розглядатися комбатантом в рамках збройних конфліктів. Проте здатність мислити, обирати та діяти як людина не надає штучно-

му інтелекту можливість нести відповідальність в рамках міжнародного гуманітарного права. До того ж, на даний момент штучний розум не наділений автономною волею та має пряму залежність від конкретної людини: командира, розробника, оператора системи, виробника [6, с. 391–392]. Відтак, можливість штучним інтелектом нести відповідальність нівелюється встановленими обмеженнями на вихідному етапі застосування.

Показовим прикладом зміни підходів провідних держав світу до штучного інтелекту є створення у лютому 2022 року Міністерством оборони США Головного управління цифрових технологій і штучного інтелекту (CDAO). Головною метою вказаного управління стало впровадження у військовий сектор штучного інтелекту, який здійснював аналіз стратегічних рішень командування для забезпечення мінімальних ризиків та втрат [7, с. 79]. У цьому розумінні, хід майбутніх воєнних дій буде зводитися до здатності автономної системи до аналізу великої кількості вхідної інформації, де дані виступають стратегічним ресурсом.

В свою чергу в Україні, державі, яка тривалий час є стороною збройного конфлікту, де застосовуються гібридні методи ведення війни, у 2020 році приймається Концепція розвитку штучного інтелекту, що ознаменувала початок роботи над регулятивними приписами, в тому числі у сфері оборони [8]. У нормативному акті, серед іншого, виокремлюються питання правосуб'єктності штучного інтелекту, а також легалізація машинного навчання в оборонних системах аналізу даних, кібербезпеки, командування та управління. Серед важливого також виокремлюється задекларована необхідність у стандартизації та сертифікації систем зі штучним інтелектом, що встановить пряму відповідальність уповноважених осіб. До того ж, будь-яке використання штучного інтелекту у військовій сфері буде проходити перевірку на відповідність ключовим міжнародно-правовим актам з прав людини.

Одним з перших імперативних нормативно-правових актів у сфері штучного інтелекту поряд з правами людини, демократією і верховенством права є Рамкова конвенція Ради Європи про штучний інтелект, права людини, демократію та верховенство права, що закріпила загальні принципи використання вказаних технологій у державному та приватному секторі, однак з винятком щодо національної оборони і безпеки, а також наукових досліджень [9, с. 1–4]. Стаття 3 вказаної конвенції містить застереження щодо можливості застосування державами штучного інтелекту у сфері національної безпеки із врахуванням приписів міжнародного права. Тобто, джерела права війни виступають основою регулювання правовідносин щодо автономної зброї під час збройних конфліктів держав.

Аналізуючи усталені міжнародні моделі відповідальності за використання штучного інтелекту під час збройних конфліктів, можна виділити наступні: «human-in-the-loop» (система не зробить постріл без свідомого натискання кнопки людиною), «human-on-the-loop» (система знаходить цілі та може робити постріли, проте людина може втрутитися) і «human-out-of-the-loop» (система наділена повною автономією) [10, с. 27]. В такому розумінні лишається відкритим питання передбачуваності конкретного військового маневру із використанням штучного інтелекту, адже будь-яке воєнне завдання передбачає конкретний результат із прорахунком можливих наслідків. В разі надання автономії штучному розуму командувач ділянки позбавляється можливості планувати військові операції.

Формальна відповідність застосування штучного інтелекту у збройних конфліктах Праву Женеви не вирішує проблему правозастосування, зокрема, встановлення відповідального суб'єкта за вчинення воєнного злочину. Оскільки штучний інтелект, як було зазначено вище, не володіє свідомістю, він не може мати умислу на вчинення воєнного злочину. Така варіативність може бути лише тоді, коли першочергово у систему був завантажений алгоритм дій, який міг передбачати протиправність дій. Відтак, у разі помилкового вибору цивільного об'єкту як цілі, постає питання: хто буде нести відповідальність? В національному кримінальному праві, як і у міжнародному (Римський статут МКС), передбачено концепцію *mens rea* (суб'єктивна сторона), яка передбачає наявність вини, мотивів та мети. Якщо мета може бути задана попередньо, то вина (умисел чи необережність) та мотив у автономної системи відсутні.

У цьому контексті концепція «human-out-of-the-loop» створює прогалини у відповідальності, оскільки вона виникає лише тоді, коли оператор або командир дотримався всіх інструкцій, а штучний інтелект все одно зробив непередбачувано і вчинив протиправні дії [11, с. 36]. В цьому випадку автономність не є гарантованою, оскільки перед наданням системі свободи, людина має закласти певну задачу та перевірити справність системи. В той же час, якщо оператор або командир, що здійснював попереднє налаштування штучного інтелекту допустив неточності або мав

певний умисел – прогалина у відповідальності зникає. Відтак, повноцінне забезпечення моделі «human-out-of-the-loop» в рамках інституту відповідальності є неможливим.

Важливим доказовим джерелом в контексті непередбачуваності автономних алгоритмів став звіт Генерального секретаря ООН А/79/224 від 23 липня 2024 року. Враховуючи, що штучний інтелект є сукупністю програмного забезпечення, що було запрограмоване на вивчення закономірностей поведінки людського розуму, відсутні гарантії безпеки від неправомірного втручання або непередбачуваних рішень. Під час криз та політичних загострень штучний інтелект може представляти загрозу ядерного масштабу [12, с. 18]. Вказане підтверджує, що світове співтовариство не впевнене в застосуванні автономних машин в збройних силах держав, оскільки справжній контроль може бути досягнений лише прямим обмеженням. Автономність штучного інтелекту в цьому розумінні не може бути досягнене через передачу контролю над бойовими частинами військового озброєння.

Пізніше Генеральною Асамблеєю ООН була прийнята резолюція А/79/239 від 24 грудня 2024 року, що ознаменувала початок збору інформації серед держав-членів щодо можливостей і викликів використання штучного інтелекту у військовій сфері, закріпила твердження про застосовність міжнародного гуманітарного права до штучного інтелекту на всіх етапах його застосування і визначила алгоритмічне мислення як допоміжний функціонал в житті людини [13, с. 1-4]. Очевидно, що людиноцентризм залишається пріоритетним напрямом в співпраці держав-членів, а виокремлення штучного інтелекту як можливу загрозу для цивільного населення ставить під сумнів надання автономних прав системам, що здійснюють воєнні операції без втручання людини.

Під впливом політичних і правових змін у міжнародному співтоваристві приватні компанії, що є розробниками та постачальниками послуг нейромереж, почали змінювати своє ставлення до використання їхніх чат-ботів та алгоритмів у військовій сфері. Так згідно Політики користування компанії OpenAI від 23 березня 2023 року заборонялося використання нейромережі для розробки зброї, військових та воєнних завдань [14]. У грудні 2024 року публікується новина про початок співпраці Anduril Industries та OpenAI з метою розвитку і впровадження штучного інтелекту у сферу національної оборони та безпеки [15]. Хоча наразі лишається застереження про заборону завдання шкоди особі, політика користування не містить прямої заборони щодо використання нейромережі у воєнній промисловості. Відтак, фактичне використання може допускатися, але безпосередня відповідальність покладається на суб'єкта, що використовує штучний інтелект у власних інтересах.

Висновки. Загалом аналізуючи використання штучного інтелекту під час збройних конфліктів, варто наголосити, що сучасна міжнародно-правова система має достатньо регулятивних норм, що обмежують держави у застосуванні автономних систем. Право Женеві, що створювалося під впливом двох найбільш кривавих війн, передбачає достатню кількість запобіжних заходів для сторін конфлікту щодо використання новітніх технологій. Хоча принцип розрізнення і пропорційності закріплюють справедливі умови ведення бойових дій, штучний інтелект не є тим суб'єктом, який здатний в повній мірі дотриматися їх.

Сама ж правосуб'єктність штучного інтелекту не може розглядатися в контексті притягнення до відповідальності, з огляду на відсутність можливостей автономної системи до розумових здібностей, що формуються у людині під впливом антропологічних, соціальних, культурних моральних, релігійних чинників. Обмеження в цьому напрямку має забезпечуватися шляхом використання штучного інтелекту лише в рамках обробки інформації та аналізу поля бою. Автономна бойова система не здатна нести відповідальність, однак людина, будь то командир, розробник, оператор, виробник або сама держава, яка взяла на озброєння такі механізми несе пряму відповідальність перед міжнародним співтовариством за використання штучного інтелекту.

Найбільш прийнятною концепцією є «human-in-the-loop», де система передбачає лише аналіз вхідних даних та відсутність прийняття рішення без підтвердження людиною. Таким чином, виділення штучного інтелекту як окремого суб'єкта не буде потребуватися, а відповідальність за неправомірні дії буде покладатися безпосередньо на учасників збройного конфлікту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Синиця А. Проблема розуміння в контексті критики комп'ютаціоналізму. *Наукові записки Національного університету «Острозька академія». Серія «Філософія»*. Острог, 2013. Вип. 14. С. 65–69. URL: <https://eprints.oa.edu.ua/2418>.

2. McCarthy J. Programs with Common Sense. *Mechanisation of Thought Processes* / J. McCarthy. Proceedings of the Symposium of the National Physics Laboratory, London, U.K., 1956. P. 77–84.
3. Ганенко І.С., Шут В.О., Яланська М.І. Женевські конвенції та новітні технології війни: проблеми ефективного правового застосування. *Науковий вісник Ужгородського національного університету. Серія: Право*. 2025. Вип. 87. Ч. 4. С. 222–228. URL: <https://doi.org/10.24144/2307-3322.2025.87.4.35>.
4. Додатковий протокол до Женевських конвенцій від 12 серпня 1949 року, що стосується захисту жертв міжнародних збройних конфліктів (Протокол I), від 8 червня 1977 року. URL: https://zakon.rada.gov.ua/laws/show/995_199.
5. Татаренко Г.В., Карпенко М.В. Міжнародно-правове регулювання технологій штучного інтелекту в бойових діях. *Актуальні проблеми права: теорія і практика*. 2023. № 2 (46). С. 135–145. URL: <https://doi.org/10.33216/2218-5461/2023-46-2-135-145>.
6. Каряка Р.О. Міжнародно-правові засади правосуб'єктності штучного інтелекту у контексті автономних систем озброєнь. *Молодіжна наука заради миру та розвитку: матеріали IV Міжнар. наук.-практ. конф. (Чернівці, 28–29 листопада 2025 р.)*. Чернівці: Чернівецький нац. ун-т ім. Ю. Федьковича, 2025. С. 391–393.
7. Kollmann T., Kollmann K., Kollmann N. Artificial Leadership: Digital Transformation as a Leadership Task between the Chief Digital Officer and Artificial Intelligence. *International Journal of Business Science and Applied Management*. 2023. Vol. 18. Iss. 1. P. 1–15. URL: https://www.business-and-management.org/library/2023/18_1--76-95-Kollmann,%20Kollmann.pdf.
8. Про схвалення Концепції розвитку штучного інтелекту в Україні: розпорядження Кабінету Міністрів України від 02.12.2020 № 1556-р. URL: <https://zakon.rada.gov.ua/laws/show/1556-2020-%D1%8>.
9. Framework Convention on artificial intelligence and human rights, democracy and the rule of law: Council of Europe Treaty Series No. 225 / Council of Europe. Vilnius, 2024. 17 May. URL: <https://rm.coe.int/1680afae3>.
10. Gevorgyan M. Crossing Boundaries: Autonomous Weapon Systems and the Challenge of IHL Compliance. Geneva: Geneva Academy of International Humanitarian Law and Human Rights, 2021. 32 p. URL: <https://ijlet.org/wp-content/uploads/2025/01/3.3.2.pdf>.
11. Königs P. Artificial intelligence and responsibility gaps: what is the problem? *Ethics Inf Technol*. 2022. Vol. 24. No. 36. URL: <https://doi.org/10.1007/s10676-022-09643-0>.
12. Current developments in science and technology and their potential impact on international security and disarmament efforts: Report of the Secretary-General: A/79/224 / United Nations General Assembly. 2024. 23 July. 19 p. URL: <https://undocs.org/A/79/224>.
13. Artificial intelligence in the military domain and its implications for international peace and security: resolution adopted by the General Assembly on 24 December 2024: A/RES/79/239 / United Nations. 2024. 4 p. URL: <https://docs.un.org/en/a/res/79/239>.
14. Usage policies / OpenAI. 2024. 9 Jan. URL: <https://web.archive.org/web/20240109122522/https://openai.com/policies/usage-policies>.
15. Anduril Partners with OpenAI to Advance U.S. Artificial Intelligence Leadership and Protect U.S. and Allied Forces / Anduril Industries. 2024. 3 Dec. URL: <https://www.anduril.com/news/anduril-partners-with-openai-to-advance-u-s-artificial-intelligence-leadership-and-protect-u-s>.

Дата першого надходження рукопису до видання: 4.02.2026

Дата прийняття до друку рукопису після рецензування: 20.02.2026

Дата публікації: 5.03.2026

© Каряка Р.О., 2026

Стаття поширюється на умовах ліцензії CC BY 4.0