

КЛИКОВ ОЛЕКСАНДР ІГОРОВИЧ

аспірант кафедри філософії Національного юридичного університету імені Ярослава Мудрого
ORCID 0009-0005-1366-2374

УДК 340.12:[004.8: 34.023]

DOI 10.37772/2518-1718-2025-3(51)-11

ПРАВОВА ТА МОРАЛЬНА ВІДПОВІДАЛЬНІСТЬ В УМОВАХ РОЗВИТКУ АВТОНОМНИХ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

У статті обґрунтовується, що зростаюча залежність суспільства від автономних систем штучного інтелекту гостро порушує питання щодо відповідальності. Зазначено, що є низка етичних дилем, пов'язаних з розробкою та використанням штучного інтелекту. Наголошується, що в наукових колах навколо питання щодо відповідальності точаться дискусії та дебати, які демонструють величезні відмінності у можливих підходах до цього питання. Продемонстровано, що питання щодо відповідальності в контексті систем штучного інтелекту вимагає міждисциплінарного підходу. З'ясовано, що застосування філософського підходу для визначення проблем відповідальності в автономних системах сприяє забезпеченню надійності автономних систем. Охарактеризовані деякі з найпоширеніших загальних підходів до дослідження поняття відповідальності. Зосереджено увагу на доцільності соціо-технічного підходу до питання відповідальності. Наголошено, що не існує єдиного доктринального підходу до розуміння питання юридичної відповідальності, а є, як мінімум, сім основних концепцій розуміння юридичної відповідальності як правового явища. Виявлено, що різні виміри відповідальності пов'язані з міждисциплінарними викликами в проектуванні, розробці та використанні надійних автономних систем та можуть їх вирішувати. Тому забезпечення надійності автономних систем є важливим міждисциплінарним завданням. Припущено, що якщо автономні системи у своєму автоматизованому процесі прийняття рішень будуть зосереджені виключно на максимізації показників ефективності, то вони, швидше за все будуть ігнорувати питання відповідальності. Висловлено, що для того, щоб мінімізувати шкідливі наслідки та забезпечити етичну та надійну поведінку автономних систем на основі штучного інтелекту у перспективі слід дозволити їм міркувати про потенційні проблеми, пов'язані з відповідальністю.

Ключові слова: моральний статус штучного інтелекту; штучний інтелект і право; відповідальність автономних систем; кількісна оцінка ступенів відповідальності; соціо-технічний підхід до відповідальності; юридична відповідальність в технологічному контексті; відповідальність смертоносних автономних систем озброєнь.

Постановка проблеми. В умовах розвитку цифрового суспільстві виникла й дедалі актуалізується ціла низка питань стосовно розробки та використання штучного інтелекту в контексті моралі. Впровадження систем штучного інтелекту зростає в усіх суспільних сферах, що також означає, що виникають питання щодо відповідальності в різних контекстах. Тому є потреба дослідити, чи використання автономних систем подекуди призводить до проблем, що пов'язані з відповідальністю і якою мірою має сенс покладати відповідальність на штучні автономні системи.

Формулювання цілей. Метою дослідження є визначення підходів до питання відповідальності в контексті розробки та використання штучних автономних систем, що сприяє забезпеченню їх етичної та надійної поведінки.

Аналіз останніх досліджень і публікацій. Вивчення проблем правової та моральної відповідаль-

ності в технологічному контексті є доволі складним. Це пов'язане з тим, що ані в науковій літературі, ані в законодавстві немає єдиного погляду на низку спірних питань, що виникають при дослідженні цієї проблематики. Окремим аспектам юридичної відповідальності присвячені роботи відомих українських правників, зокрема питанням дослідження юридичної відповідальності як системного явища (Р. Майданик), аналізу доктринальних підходів до розуміння сутності юридичної відповідальності (В. Жорнокуй), розгляду юридичної відповідальності як форми державного примусу (В. Надьон, Р. Шишка) та ін. Хоча існує сім основних концепцій розуміння юридичної відповідальності як правового явища, проте жодна з них не дає відповіді на питання, які пов'язані з розробкою та використанням штучних інтелектуальних автономних систем. Дослідженням окремих правових питань теми технологій присвячені роботи ві-

тизних науковців Д. Адамюка, Г. Андрощука, О. Давидюка, Б. Падучака, М. Швеця та ін. У вітчизняній правовій науці питання моральної відповідальності в технологічному контексті майже не досліджувалося. Аналіз проблем правової та моральної відповідальності штучних автономних систем вимагає нових методологічних схем.

Виклад основного матеріалу. Постійно зростаюча залежність суспільства від штучних автономних систем гостро порушує питання щодо відповідальності. Треба зазначити, що більшість дослідників вважають, що на всіх етапах проектування, розробки та розгортання надійних автономних систем міркування про відповідальність повинні відігравати ключову роль.

В наукових колах в технологічно розвинених країнах сьогодні активно обговорюють питання щодо відповідальності та прогалин у відповідальності в контексті систем штучного інтелекту (тобто ситуацій, коли група несе відповідальність, але відповідальність окремих осіб може бути неясною) і які наслідки з ними пов'язані, а також що потрібно зробити для усунення потенційних прогалин у відповідальності. Наприклад, коли автономні системи роблять помилку, то хто має нести за це відповідальність?

В країнах Євросоюзу підставою для дискусії про відповідальність автономних систем постали положення Резолюції щодо норм цивільного права про робототехніку від 16.02.2017 р. (2015/2013(INL)), що була прийнята Європейським парламентом. Цим документом передбачалося положення щодо можливості надання автономним роботам статусу електронної особи для врегулювання питання відповідальності за заподіяну ними шкоду, а також, застосування цього поняття до випадків, коли роботи приймають автономні рішення або незалежно (самостійно) вступають у відносини з третіми особами.

Сьогодні поширена думка, що відповідальність за помилки автономних систем має бути покладена на розробників або провайдерів, оскільки саме люди повинні контролювати відповідні рішення, а отже, нести за них відповідальність. У Регламенті ЄС, що встановлює гармонізовані правила щодо штучного інтелекту 2024/1689 від 13 червня 2024 року був закладений підхід, що відповідальність за помилки автономних систем має бути покладена на розробників або провайдерів [1].

З точки зору низки провідних українських науковців автономні системи варто розглядати як об'єкт правовідносин (О.А. Баранов, Є.О. Мічурін, В.Л. Яроцький та інші). Українське законодавство на сьогодні не містить норм, які б визнавали автономні системи суб'єктом правовідносин. Тому, доки такі системи визначають як об'єкт правовідносин, відпо-

відальність покладатиметься на розробника або користувача, однак у зв'язку з подальшим стрімким розвитком технологій штучного інтелекту не виключено, що незабаром ситуація зміниться. Окрім цього, практичний досвід використання автономних систем на прикладі автономних транспортних засобів показує, що люди не завжди здатні розуміти, яку частину штучної системи вони повинні взяти під контроль і в який відповідний момент. Усвідомлення цього аспекту показує необхідність міждисциплінарного дослідження питання про відповідальність, щоб вирішити, хто і якою мірою відповідає за довгострокову поведінку системи штучного інтелекту або за конкретне рішення, прийняте нею.

Деякі з відомих філософів, зокрема Андреас Маттіас та Роберт Спарроу вважають, що невинувато притягувати когось до відповідальності за помилки автономних систем. Якщо, наприклад, безпілотний автомобіль збиває когось, на їхню думку, ніхто не може бути притягнутий до відповідальності [2].

На думку низки дослідників, одним із найсуперечливіших прикладів є розвиток смертоносних автономних систем озброєння, які в наукових джерелах отримали назву «роботів-вбивць» або «автономних систем озброєння». Треба зазначити, що дотепер немає ані узгодженого на міжнародному рівні визначення поняття «автономна система озброєння» чи «смертоносна автономна система озброєння», ані загального опису його характеристик, ані навіть загальноприйнятого іменування такого поняття [3, с. 243.], хоча в резолюціях ООН прийнято використовувати термін «смертоносні автономні системи озброєння» [4]. Очевидним фактом є той, що розвиток цих систем озброєння докорінно змінить природу конфліктів і таким чином, створить фундаментальні проблеми для стабільності міжнародної системи, оскільки такі системи самостійно обирають цілі й приймають рішення про атаку без участі людини. На сьогодні до кінця ще не зрозуміло, якою мірою така технологія в повністю автономній формі та її використання в смертоносних цілях у війні вже існує [2], проте зрозуміло, що використання штучного інтелекту у війні швидко зростає. Війна в Україні й Ізраїльський конфлікт в секторі Газа прискорили розробки автономних систем озброєння, які постають як ключові технології майбутніх воєнних конфліктів [5]. Наприклад, згідно з інформацією з відкритих джерел [6], український комплекс з можливостями штучного інтелекту для автономних місій Saker Scout може автономно літати над визначеними територіями без контакту з людиною, уникаючи проблем, пов'язаних з глушінням GPS або електронними перешкодами та здатний розпізнавати і фіксувати цілі та наводити на них інших дронів, може виступати та-

кож як камікадзе, діючи на останньому етапі, після пригнічення каналу зв'язку системами РЕБ ворога, в автономному режимі та вражаючи ціль уже без участі оператора. Ізраїльська система баражуючого боєприпасу зі штучним інтелектом Lanius може автономно виконувати повний профіль польоту, зліт, навігацію і розвідку без втручання користувача, а у таких місцях, як двері та кути, може використовувати режим засідки для швидкої атаки у відповідь на ознаки руху, а також може автономно ідентифікувати точки виходу і входу, людські цілі за допомогою відеоаналізу на основі штучного інтелекту і присвоювати об'єктам і людям класифікацію загрози [6]. Хоча зараз дронами у повітрі, воді та на суші керують люди, які й приймають рішення щодо обрання воєнних цілей, окремі зразки роботизованих артефактів у своєму розвитку вже наблизилися до ознак так званих «роботів-убивць» [6].

Питання щодо відповідальності в контексті систем штучного інтелекту є дуже складним і багатограним. Будь-які взаємовідносини між людьми, їх вчинки характеризуються відповідальністю, вона стосується майже всіх аспектів людської діяльності, тому «відповідальність» не є юридичним терміном. Поняття «юридична відповідальність» знаходиться у відношенні підпорядкування до поняття «відповідальність», яке виражає свідоме ставлення особи до суспільної необхідності, загальних потреб, обов'язків, норм, цінностей [7]. Відповідальність, як така, є соціологічною категорією, її навіть часто узагальнюють в понятті «соціальна відповідальність», розглядають як підзвітність (accountability) суб'єкта за свої дії і вчинки [8, р. 259], або моральний обов'язок, що пов'язаний з виконанням тих чи інших покладених на окрему особу або групу осіб зобов'язань [9, р. 839] або здатність суб'єкта передбачати результат своїх дій, як почуття службового або іншого обов'язку й усвідомлення своєї відповідальності [10, с. 63].

В наукових працях поняття «відповідальність» досліджують в різних аспектах. Період з кінця XIX ст. і дотепер характеризується утвердженням різних точок зору щодо поняття відповідальності. Сьогодні найпоширенішими серед філософських підходів до дослідження поняття відповідальності є 1) етична концепція М. Вебера (послідовниками якої є Г. Йонас, М. Рідель, Ю. Габермас, К.-О. Апел), яка розглядає відповідальність як етичну категорію, в якій «перехрещуються мовні та реальні відносини особи, що виражається у відповідності з усіма членами комунікативної спільноти» [11, с. 81]; 2) інституційний підхід, представниками якого є Т. Парсонс, Е. Дюркгейм, М. Вебер, В. Ядов, Дж. Ролз, П. Рікер та ін., у межах якого ідеологічність має обов'язковий характер щодо діяльності всіх соціальних інститутів

та 3) екзистенціальний підхід, відповідно до якого відповідальність розглядається як характеристика людини.

Психологи визначають відповідальність як «внутрішню саморегуляцію і самодетермінацію зрілої особистості, що опосередкована ціннісними орієнтаціями і виявляється в усвідомленні людиною причин здійснюваних учинків і їх наслідків, у контролі своєї здатності бути причиною змін у навколишньому світі і власному житті» [12, с. 60]. Американський психолог Карл Роджерс розглядає відповідальність як основну характеристику сформованої особистості. В роботі американсько-німецького науковця Г. Йонаса «Принцип відповідальності. У пошуках етики для технологічної цивілізації» в рамках етичної концепції, акцентується увага на техногенній небезпеці, що загрожує суспільству [13].

Взагалі розрізняють причинно-наслідкову відповідальність, моральну відповідальність та рольову відповідальність, які пов'язані між собою. Моральна відповідальність завжди ґрунтується на причинно-наслідковій, але причинно-наслідкова відповідальність може виникати у людини, яка не є відповідальною в моральному сенсі.

В юридичній науці питання відповідальності є одним з найчастіше обговорюваних і вважається доволі складним, оскільки не існує єдиного доктринального підходу до його розуміння. Існує, як мінімум, сім основних концепцій розуміння юридичної відповідальності як правового явища [14, с. 91], причому жодна з них не дає однозначної відповіді на поставлене питання. На думку проф. В. Г. Жорнокуй конструювання загальноновизнаного визначення поняття юридичної відповідальності сьогодні є недоцільним, оскільки «в багатьох випадках фахівці прагнуть об'єднати численні та різноманітні ознаки юридичної відповідальності у визначення, що залишає поза увагою специфічні прояви різних аспектів її феномена. Наявні сьогодні доктринальні підходи до розуміння юридичної відповідальності можна звести до двох основних напрямів її вивчення як охоронних правовідносин або форми державного примусу/впливу» [14, с. 97].

Висновки. Оскільки дедалі стає зрозуміло, що неможливо очікувати, щоб всі компоненти автономної системи поведилися належним чином, то виникає потреба в тому, щоб розробити й запровадити загальні методи для забезпечення надійності та стійкості. Поняття відповідальності може бути основою для концептуалізації надійності та стійкості, тому виникає необхідність розробити комплекс методів, здатних аналізувати компроміс між ефективністю та стійкістю в соціо-технічних системах та розробляти комплексні моделі стійкого партнерства між людиною та штучним інтелектом.

Забезпечення надійності автономних систем та штучного інтелекту є важливим міждисциплінарним завданням, до розв'язання якого, зокрема, може бути застосований філософський підхід для визначення проблем відповідальності в автономних системах, соціо-технічний підхід для кількісної оцінки ступенів відповідальності та координації завдань на основі цього [5] та правовий підхід для визначення поняття юридичної відповідальності в технологічному контексті.

Якщо у перспективі дозволити автономним системам на основі штучного інтелекту міркувати про потенційні проблеми, пов'язані з відповідальністю,

то можна мінімізувати шкідливі наслідки та забезпечити етичну та надійну поведінку. Якщо автономні системи у своєму автоматизованому процесі прийняття рішень будуть зосереджені виключно на максимізації показників ефективності, то вони, швидше за все будуть ігнорувати питання відповідальності (від підзвітності до юридичної відповідальності). Тому сьогодні актуальним є комплексне дослідження даного питання в технологічному контексті, щоб з'ясувати, хто і якою мірою відповідає за довгострокову поведінку автономної системи штучного інтелекту або за конкретне рішення, прийняте нею.

ЛІТЕРАТУРА

1. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). *EUR-Lex Access to European Union law*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
2. Yazdanpanah, V., Gerding, E.H., Stein, S. et al. Reasoning about responsibility in autonomous systems: challenges and opportunities. *AI & Soc* 38, 1453–1464 (2023). <https://doi.org/10.1007/s00146-022-01607-8>
3. Суспільство, людина, право: сучасний соціокультурний контекст : монографія / [О. П. Андрущенко, О. Г. Данильян, О. П. Дзьобань та ін.] ; за ред. О. Г. Данильяна. ; Нац. юрид. ун-т ім. Ярослава Мудрого ; НДІ держ. буд-ва та місц. самоврядування НАПРН України. Харків : Право, 2025. 328 с.
4. Lethal autonomous weapons systems : report of the Secretary-General. UN Digital Library. 2024. URL: <https://digitallibrary.un.org/record/4059475?ln=en&v=pdf>.
5. Pedro A., Martínez M. Los Sistemas de Armas Autónomos Letales y el Derecho Internacional Humanitario en la Guerra de Ucrania. *Relaciones Internacionales*. 2023. N 53. P. 71-90. DOI: <https://doi.org/10.15366/relacionesinternacionales2023.53.004>
6. Короткий В. Зброя і штучний інтелект: у світі занепокоєні збереженням людського контролю. *Укрінформ*. 06.05.2024. URL: <https://www.ukrinform.ua/rubric-world/3860523-zbroia-i-stucnij-intelekt-u-sviti-zanepokoeni-zberezennam-ludskogo-kontrolyu.html> (дата звернення: 13.02.2025)
7. Енциклопедія освіти / Акад. пед. наук України ; голов. ред. В. Г. Кремень. Київ : Юрінком Інтер, 2008. 1040 с.
8. Fairchild H. P. Dictionary of Sociology and Related Sciences. Littlefield : Adams, 1960. 342 p.
9. Hornby A. S. Oxford Advanced Learner's Dictionary of Current English. Oxford : Oxford University Press, 1998. 1476 p.
10. Марченко М. Проблеми юридичної та соціально-політичної відповідальності бізнесу. *Про українське право*. 2010. № 5. С. 62–69.
11. Вітковська І. М., Євдокимова І. А. Становлення поняття відповідальності у філософії та соціології. *Актуальні проблеми філософії та соціології*. 2020. С. 77-81. DOI <https://doi.org/10.32837/apfs.v0i27.925>
12. Психологічна енциклопедія / авт.-упоряд. О. М. Степанов. Київ : Академвидав, 2006. 424 с.
13. Йонас Г. Принцип відповідальності. У пошуках етики для технологічної цивілізації. Київ : Лібра, 2001. 400 с.
14. Жорнокуй В. Г. Юридична відповідальність: сім доктринальних підходів до розуміння сутності. *Право і безпека*. 2024. № 1 (92). С. 90–100. DOI: <https://doi.org/10.32631/pb.2024.1.08>.

REFERENCES

1. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). *EUR-Lex Access to European Union law*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng> [in English].

2. Yazdanpanah, V., Gerding, E.H., Stein, S. et al. (2023). Reasoning about responsibility in autonomous systems: challenges and opportunities. *AI & Soc* 38, 1453–1464 <https://doi.org/10.1007/s00146-022-01607-8> [in English].
3. Suspilstvo, liudyna, pravo: suchasnyi sotsiokulturnyi kontekst : monohrafiia (2025) / [O. P. Andrushchenko, O. H. Danylian, O. P. Dzoban ta in.] ; za red. O. H. Danyliana. ; Nats. yuryd. un-t im. Yaroslava Mudroho ; NDI derzh. bud-va ta mists. samovriaduvannia NAPrN Ukrainy. Kharkiv : Pravo, 2025. 328 s. [in Ukrainian].
4. Lethal autonomous weapons systems : report of the Secretary-General. UN Digital Library. 2024. URL: <https://digitallibrary.un.org/record/4059475?ln=en&v=pdf>. [in English].
5. Pedro A., Martínez M. (2023). Los Sistemas de Armas Autónomos Letales y el Derecho Internacional Humanitario en la Guerra de Ucrania. *Relaciones Internacionales*. N 53. P. 71-90. DOI: <https://doi.org/10.15366/relacionesinternacionales2023.53.004> [in Spanish].
6. Korotkyi V. (2024). Zbroia i shtuchnyi intelekt: u sviti zanepokoieni zberezheniam liudskoho kontroliu. *Ukrinform*. 06.05.2024. URL: <https://www.ukrinform.ua/rubric-world/3860523-zbroia-i-stucnij-intelekt-u-sviti-zanepokoieni-zberezheniam-ludskogo-kontroliu.html>
7. Entsiklopediia osvity (2008). Akad. ped. nauk Ukrainy ; holov. red. V. H. Kremen. Kyiv : Yurinkom Inter. 1040 s. [in Ukrainian].
8. Fairchild H. P. (1960). Dictionary of Sociology and Related Sciences. Littlefield : Adams, 342 p. [in English].
9. Hornby A. S. (1998). Oxford Advanced Learner's Dictionary of Current English. Oxford : Oxford University Press, 1476 p. [in English].
10. Marchenko M. (2010). Problemy yurydychnoi ta sotsialno-politychnoi vidpovidalnosti biznesu. *Pro ukrainske pravo*. № 5. S. 62–69. [in Ukrainian].
11. Vitkovska I. M., Yevdokymova I. A. (2020). Stanovlennia poniattia vidpovidalnosti u filosofii ta sotsiologii. *Aktualni problemy filosofii ta sotsiologii*. S. 77-81. DOI <https://doi.org/10.32837/apfs.v0i27.925> [in Ukrainian].
12. Psykholohichna entsiklopediia / avt.-uporiad. O. M. Stepanov. (2006). Kyiv : Akademvydav, 424 s. [in Ukrainian].
13. Yonas H. (2001). Pryntsyp vidpovidalnosti. U poshukakh etyky dlia tekhnolohichnoi tsyvilizatsii. Kyiv : Libra, 400 s. [in Ukrainian].
14. Zhornokui V. H. (2024). Yurydychna vidpovidalnist: sim doktrynalnykh pidkhodiv do rozuminnia sutnosti. *Pravo i bezpeka*. 2024. № 1 (92). S. 90–100. DOI: <https://doi.org/10.32631/pb.2024.1.08>. [in Ukrainian].

KLYKOV OLEKSANDR

Graduate Student of the Department of Philosophy of the Yaroslav Mudryi National Law University

LEGAL AND MORAL RESPONSIBILITY IN THE CONDITIONS OF DEVELOPMENT OF AUTONOMOUS ARTIFICIAL INTELLIGENCE SYSTEMS

Problem setting. In the context of the development of a digital society, a number of increasingly relevant questions have arisen regarding the development and use of artificial intelligence in relation to morality. The implementation of artificial intelligence systems is expanding in all spheres of society, which also means that questions of responsibility arise in various contexts. Therefore, it is necessary to investigate whether the use of autonomous systems sometimes leads to responsibility-related problems and to what extent it makes sense to assign responsibility to artificial autonomous systems.

Analysis of recent researches and publications. The study of issues of legal and moral responsibility in a technological context is quite complex. This is due to the fact that neither academic literature nor legislation provides a unified view on a number of contentious questions that arise when examining this topic. Certain aspects of legal responsibility have been explored in the works of well-known Ukrainian legal scholars, including studies on legal responsibility as a systemic phenomenon (R. Maidanyk), analyses of doctrinal approaches to understanding the essence of legal responsibility (V. Zhornokui), and examinations of legal responsibility as a form of state coercion (V. Nadeon, R. Shyshka), among others. Although there are seven main concepts for understanding legal responsibility as a legal phenomenon, none of them provides answers to the questions associated with the development and use of artificial autonomous intelligent systems. Certain legal aspects related to technology have been addressed in the works of Ukrainian scholars such as D. Adamiuk, H. Androshchuk, O. Davydiuk, B. Paduchak, M. Shvets, and others. In Ukrainian legal scholarship, the issue of moral responsibility in the technological context has been scarcely explored. The analysis of legal and moral responsibility of artificial autonomous systems requires new methodological frameworks.

Purpose of the research. The aim of the study is to identify approaches to the issue of responsibility in the context of the development and use of artificial autonomous systems, which contributes to ensuring their ethical and reliable behavior.

Article's main body. The article argues that society's growing dependence on artificial-intelligence autonomous systems sharply raises questions of responsibility. It is noted that there are a number of ethical dilemmas associated with the development and use of artificial intelligence. The article emphasizes that discussions and debates continue within the academic community regarding the issue of responsibility, demonstrating significant differences in possible approaches to this topic. It is shown that responsibility in the context of AI systems requires an interdisciplinary approach. It is established that applying a philosophical approach to identifying responsibility-related problems in autonomous systems contributes to ensuring the reliability of such systems. Some of the most common general approaches to studying the concept of responsibility are characterized. Attention is focused on the relevance of the sociotechnical approach to the issue of responsibility. It is stressed that there is no single doctrinal approach to understanding legal responsibility; rather, there are at least seven main concepts interpreting legal responsibility as a legal phenomenon. It is identified that different dimensions of responsibility are connected to interdisciplinary challenges in the design, development, and use of reliable autonomous systems and can help address them. Therefore, ensuring the reliability of autonomous systems is an important interdisciplinary task.

It is assumed that if autonomous systems in their automated decision-making processes focus exclusively on maximizing efficiency indicators, they are likely to ignore responsibility-related issues. It is concluded that, in order to minimize harmful consequences and ensure ethical and reliable behavior of AI-based autonomous systems, they should, in the future, be allowed to reason about potential responsibility-related problems.

Conclusions and prospects for development. As it is becoming increasingly clear that it is impossible to expect all components of an autonomous system to behave properly, there is a need to develop and implement general methods to ensure reliability and resilience. The concept of responsibility can serve as a basis for conceptualizing reliability and resilience, which creates the need to develop a set of methods capable of analyzing the trade-off between efficiency and resilience in sociotechnical systems, as well as designing comprehensive models of sustainable collaboration between humans and artificial intelligence.

Ensuring the reliability of autonomous systems and artificial intelligence is an important interdisciplinary task. A philosophical approach can be applied to identify issues of responsibility in autonomous systems; a sociotechnical approach can be used to quantitatively assess degrees of responsibility and task coordination based on this; and a legal approach can be used to define the concept of legal responsibility in a technological context.

If, in the future, artificial-intelligence-based autonomous systems are allowed to reason about potential responsibility-related problems, harmful consequences can be minimized, and ethical and reliable behavior can be ensured. If autonomous systems in their automated decision-making processes focus exclusively on maximizing efficiency indicators, they are likely to ignore issues of responsibility (from accountability to legal liability). Therefore, a comprehensive study of this issue in the technological context is currently relevant in order to determine who, and to what extent, is responsible for the long-term behavior of an artificial intelligence autonomous system or for a particular decision it makes.

Key words: moral status of artificial intelligence; artificial intelligence and law; responsibility of autonomous systems; quantification of degrees of responsibility; sociotechnical approach to responsibility; legal responsibility in a technological context; responsibility of lethal autonomous weapons systems.

За ДСТУ 8302:2015 цю статтю слід цитувати:

Кликов О. І. Правова та моральна відповідальність в умовах розвитку автономних систем штучного інтелекту. *Право та інновації*. 2025. № 3 (51). С. 97–102.