

УДК 340.1

DOI [https://doi.org/10.20535/2308-5053.2022.3\(55\).269563](https://doi.org/10.20535/2308-5053.2022.3(55).269563)

ЕТИКО-ПРАВОВІ ПРОБЛЕМИ ВИКОРИСТАННЯ РОБОТІВ ЗІ ШТУЧНИМ ІНТЕЛЕКТОМ

Головко О. М.,

*кандидат юридичних наук, старший дослідник,
старший викладач кафедри інтелектуальної власності та приватного права
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID ID: 0000-0001-8963-6598*

Боднар Є.,

*випускниця курсу «Європейська інтеграція: законодавство та Інтернет речей» (EULIOT),
юрист, директор
ПКП «Консульт»
ORCID ID: 0000-0002-0680-8831*

У статті розкриваються проблеми формування морально-етичних норм, якими мають керуватись роботи зі штучним інтелектом у процесі прийняття автономних рішень. Зазначено, що досягнення людства в сучасному світі є результатом формування спільних цінностей та дотримання не лише правових норм, а й суспільної моралі. Обґрунтовано, що європейський досвід у даній сфері є цінним надбанням, яке потребує дослідження та поширення. У статті значну увагу зосереджено на тому, що процес розвитку інноваційних технологій створює нові напрями досліджень суспільних відносин, які потребують врегулювання як з правової, так і з морально-етичної точки зору. Обґрунтовано, що у довгостроковій перспективі людству необхідно буде розвивати власне бачення у напрямку дотримання етично-правових норм роботами зі штучним інтелектом.

Наявні на теперішній час правові акти наділені різними видами етичних та правових норм, що дозволяє уникнути ситуацій з порушенням прав людини. Однак, постає питання досягнення балансу між морально-етичними і правовими нормами, які існують у суспільстві в ситуації, коли роль цих норм є вирішальною не тільки в поведінці людей, а й у поведінці роботів зі штучним інтелектом.

Ключові слова: етичні норми, роботи, штучний інтелект, законодавство ЄС.

Постановка проблеми. Сучасний рівень розвитку людства формує потребу появи глобальної стратегії вирішення головної суспільної суперечності між зростаючими потребами людства і неможливістю задоволення цих потреб поточними методами і засобами. Зазначену суперечність можливо вирішити завдяки штучному інтелекту, однак коректне та безпечне використання цих технологій потребує «прищеплення» певних цінностей і моральних норм. Використання об'ємних баз даних, зокрема, з всесвітньої мережі Інтернет, можливість зберігання та обробки значних об'ємів інформації – це та цифрова віртуальна дійсність, яка є простором навчання для штучного інтелекту. Саме цей код навчального процесу вважається основою технологій штучного інтелекту. Саме він може стати способом успішного формування захисного щита морально-етичних норм в алгоритмі прийняття рішень штучним інтелектом.

На підставі цього, **метою статті** є дослідження особливостей правового забезпечення морально-етичних аспектів застосування роботів зі штучним інтелектом. Основними **завданнями** є дослідження проблемних аспектів дотримання морально-етичних норм роботами з штучним інтелектом, з'ясування основних шляхів вирішення зазначених проблем, зокрема, за допомогою права, а також занурення в європейський досвід.

Аналіз останніх досліджень і публікацій. Проблемні питання морально-етичних аспектів використання людьми роботів зі штучним інтелектом були предметом дослідження ряду науковців. Серед вчених, які досліджували окремі аспекти даної проблеми, доцільно виокремити праці Баранова О. А., Бежевець А. М., Позової Д. Д. та інших вчених.

Виклад основного матеріалу. Важливо відзначити, що мораль і право, як основні соціальні регулятори, що поєднані між собою та доповнюють один одного, є результатом творчої діяльності інтелекту людини. В результаті науковому прогресу постає проблемне питання про те, чи можуть мораль і право бути сприйняті штучним інтелектом (далі – ШІ), який є протилежним від інтелекту фізичної особи, та чи може такий інтелект керуватися ними при прийнятті автономних рішень. Та й взагалі, що вбачається як процес прийняття рішення для системи ШІ; які моральні та правові наслідки їх дій і рішень; чи можна притягнути систему ШІ до відповідальності за свої дії. Це далеко не повний перелік питань, які виникають під час підняття теми норм (як правових, так і морально-етичних), які мають регулювати питання, пов'язані з використанням технологій ШІ.

Між тим, норми моралі, на відміну від норм права, далеко не завжди мають формальне визначення, а отже, такі норми фактично не є загальноновизначеними для всіх. Крім того, одне й те саме суспільство може налічувати в собі різні, відмінні і, навіть, взаємно протилежні сукупності моралі (релігійна мораль різноманітних віросповідань, корпоративна та етнічна мораль або пережитки минулих соціальних та політичних систем, так звана «імперіалістична мораль», яка може виправдовувати злочинність своїх намірів вищим благом тощо). Оскільки мораль не має уніфікованих правил поведінки, то їх дотримання не може формувати певний правовий обов'язок будь-якому суб'єкту взаємовідносин. Дійсно, ті морально-етичні норми, які не знайшли свого відображення в правових нормах, можуть спричинити, хіба що суспільний осуд. Звичайно, їх порушення може тягнути за собою і певні репутаційні ризики. Однак, якщо мова йде про ШІ, то ні репутаційні ризики, ні громадський осуд не можуть стати запобіжниками порушення етичних норм. Тому, розподіл соціальних норм на етичні та правові не працюватиме в роз'єднаному вигляді щодо ШІ. З цієї позиції варто висловити припущення, що дані норми мають становити комплекс правил поведінки, які будуть створювати рамки для тих автономних рішень, які може приймати ШІ.

Цю проблематику дуже добре ілюструє так звана проблема вагонетки, «Trolley problem» – уявний експеримент в етиці, під час якого слід пожертвувати одним або декількома життями, щоб врятувати інше або інші (Mikhail, 2007). Людство ще не вигадало ідеального готового рішення для цієї проблеми. Однак, можна відштовхуватись, наприклад, від поняття крайньої необхідності, в змісті якого допускається спричинення меншої шкоди задля запобігання або усунення реальної небезпеки чи більшої шкоди. Чи варто передбачити такий алгоритм прийняття рішень штучним інтелектом? В такому випадку, проблема вагонетки може вирішуватись жертвуванням навіть водієм. А тепер риторичне питання: який покупець зажадає купити авто з ШІ, в алгоритмі якого допускається жертвування своїм власником заради збереження більшої кількості життів (наприклад, натовпу пішоходів)?

Особа пізнає норми моралі протягом власного життя через свій особистий досвід. Він може визначатись і відповідністю формальним ознакам, виходячи з вимог законодавства, наприклад, досягнення віку юридичної відповідальності. При цьому, даний вік визначається, виходячи з того, що дитина, яка виховується в соціумі та спирається на морально-етичний комплекс правил поведінки, по досягненні віку тієї ж кримінальної відповідальності вже усвідомлюватиме неприпустимість вчинення правопорушень. Важливо відзначити, що для штучного інтелекту не є складним вивчити, проаналізувати та систематизувати чітко визначені моральні та правові норми у більш короткий відлік часу, ніж це вимагається від людини. Тож питання про те, чи може штучний інтелект засвоїти норми моралі, або чи зможе він спиратися на ці норми у власній поведінці ймовірніше має позитивну відповідь.

Таким чином, у разі дотримання штучним інтелектом універсальних норм і приписів моралі, обов'язковим вважається їх відображення в певному зведеному акті правил поведінки, так само як здійснюється систематизація правових норм. Джоанна Брайсон з цього приводу доречно підкреслює, що місце ШІ в суспільстві є питанням нормативної, а не описової етики (Bryson, 2018).

З правової точки зору проблемні аспекти морально-етичної нерівності людей і роботів зі штучним інтелектом потребує більш кращого врегулювання на законодавчому рівні, зокрема, необхідно прийняти відповідні нормативно-правові акти, які мають деталізувати порядок використання роботів зі штучним інтелектом та їх правовий статус у суспільстві (Бежевець, 2019).

Ця деталізація дійсно є необхідною, втім така законопроектна робота потребує більш комплексного підходу. Як вказують деякі науковці, для цілей правового аналізу змісту суспільних відносин не важливо знання технічних або технологічних особливостей технічних засобів, за допомогою яких вони реалізуються. Це відповідає одному з фундаментальних принципів правової доктрини – технологічної нейтральності правового регулювання. Технічні або технологічні особливості деяких засобів,

в тому числі і ШІ, які використовуються під час реалізації суспільних відносин, можуть викликати тільки деякі особливості реалізації цих відносин, що в свою чергу зумовлює можливість появи деяких особливостей правового регулювання, яке в основі своїй залишається незмінним (Позова, 2017).

В той же час, Баранов О. А. доречно підкреслює міждисциплінарність досліджень, пов'язаних з ШІ, наголошуючи, що для формування системи правового регулювання «доцільно організувати проведення системних досліджень із залученням науковців, експертів та спеціалістів зі сфер штучного інтелекту, інформаційних технологій, робототехніки, комп'ютерної інженерії, програмування, нейробиології, нейропсихології, психології та ін» (Баранов, 2022).

Активне використання робототехніки у різних сферах життя людини зумовлює розробку не тільки тих етичних аспектів, що мають бути передбачені в алгоритмі прийняття ШІ рішень, але й етичних принципів для проектувальників, виробників та користувачів роботів. Сучасне суспільство безперечно потребує розробки теорій, методів та алгоритмів, необхідних для інтеграції суспільних, правових і моральних цінностей у технологічні розробки ШІ на всіх етапах розвитку (аналіз, проектування, створення, використання та оцінка, тобто аналіз використання технологій ШІ).

Саме тому, коли постає питання відповідальності ШІ, першочергово та беззаперечно має стояти питання відповідальності людини за розвиток таких технологій. Тому варто погодитись з тезою про те, що «відповідальний штучний інтелект – це відповідальність людини за розвиток інтелектуальних систем відповідно до фундаментальних людських принципів і цінностей, щоб забезпечити процвітання та добробут людей у сталому світі» (Dignum, 2018)

Зважаючи на це, та повертаючись до теми уніфікації норм, яких повинен дотримуватись штучний інтелект, варто підкреслити цінність напрацювань, які вже має європейська співтовариство. Так, у 2019 році Європейською комісією було опубліковано Рекомендації з етики для надійного штучного інтелекту (далі – Рекомендації) (European Commission, 2019). Дана Рекомендація задекларувала наступні принципи: поваги до автономії людини, запобігання шкоді, справедливості і пояснення. Також зазначалось про необхідність визнання та вирішення потенційних протиріч між цими принципами. Окрема увага звертається на ситуації, пов'язані з більш уразливими групами населення, зокрема, дітьми та людьми з обмеженими можливостями, а також на ситуації, які характеризуються асиметрією влади чи інформації, як-от між роботодавцями та працівниками, або між підприємствами та споживачами. Це говорить про те, що суб'єктний склад правовідносин, в яких використовується штучний інтелект є важливим фактором, який теж потрібно враховувати при формуванні етичних норм для ШІ.

Основна мета даної Рекомендації є сприяння надійному ШІ. Однак, що ж передбачає конструкт «надійний штучний інтелект»? Він є логічним продовженням вищезгаданого відповідального штучного інтелекту, відповідальність якого – це відповідальність людини, яка його розробила, технічно обслуговує чи використовує. З цієї позиції варто говорити про етичний аспект не тільки в алгоритмі ШІ, але й поведінці людини.

Отже, в Рекомендації надійний штучний інтелект має три компоненти, які повинні бути протягом усього життєвого циклу системи:

- 1) Законність;
- 2) Етичність;
- 3) Надійність.

Всі ці три компоненти є невід'ємними та взаємопроникними, адже надійність має бути як з технічної, так і з соціальної точки зору, оскільки «навіть з благими намірами системи штучного інтелекту можуть спричинити ненавмисну шкоду». Саме тут ми й повертаємось до питання відповідальності людини за результати дій ШІ, оскільки поки що немає підстав говорити про наявність у ШІ проявів волі, а отже категорії «умисел» та «необережність» не можуть бути застосовними до нього.

Висновки. Основні властивості, які мають бути притаманні ШІ – це поєднання необхідності дотримання не тільки правових, але й морально-етичних норм, які не завжди прописані в нормативно-правових актах, здатність до навчання, а також можливість аналізувати та приймати рішення відповідно до встановлених у суспільстві норм. Майбутні напрямки використання ШІ є істотно важливими: алгоритми, що надають можливість щогодини обробляти великі об'єми інформації, зможуть виявляти причинно-наслідкові зв'язки, встановлення яких у короткі проміжки часу не під силу людині, а відповідно, зробити прогнози точнішими, а рішення більш раціональними. Вирішення головної суспільної суперечності розвитку між зростаючими потребами людства і неможливістю задоволення усіх потреб суспільства потребує нових рішень, одним з яких і є використання ШІ. Зазначену суперечність можливо вирішити лише за допомогою пошуку аксіологічної раціональності суспільного функціонування, що передбачає раціональне, а не хаотичне використання ШІ.

Це потребує додаткової прогностичної оцінки в правовому та морально-етичному аспектах, що унеможливить перебування людства в невідповідному стані у разі виникнення нових викликів

пов'язаних з використанням ШІ. Рекомендації Єврокомісії базуються саме на цьому. Варто пам'ятати, що мета у вигляді надійного штучного інтелекту може бути досягнута лише тоді, коли людина візьме на себе відповідальність щодо формування такої надійності на різних етапах розробки, обігу та використання технологій ШІ.

Дослідження проведено в межах курсу «European Integration: legislation and the IoT» («Європейська інтеграція: законодавство та Інтернет речей») міжнародного проекту в сфері освіти напрому Жан Моне «Модуль» програми «Erasmus+» № 620017-EPP-1-2020-1-UA-EPPJMO-MODULE (спільний проект КПІ ім. Ігоря Сікорського, Еразмус+ Жан Моне Фонду та Виконавчого агентства з питань освіти, аудіовізуальної діяльності та культури за підтримки ЄС)». Підтримка Європейською комісією випуску цієї публікації не означає схвалення змісту, який відображає лише думки авторів, і Комісія не може нести відповідальність за будь-яке використання інформації, що міститься в ній.

Golovko O., Bodnar E. Ethical and legal problems of using robots with artificial intelligence

The article reveals the problems of forming moral and ethical norms that should guide robots with artificial intelligence in the process of making autonomous decisions. It is noted that the achievements of mankind in the modern world are the result of the formation of common values and observance of not only legal norms, but also social morality. It is substantiated that the European experience in this field is a valuable asset that needs research and dissemination. In the article, considerable attention is focused on the fact that the process of development of innovative technologies creates new directions of research into social relations, which need to be regulated from both a legal and a moral and ethical point of view. It is substantiated that in the long term, humanity will need to develop its own vision in the direction of compliance with ethical and legal norms by robots with artificial intelligence.

Currently available legal acts are endowed with various types of ethical and legal norms, which allows to avoid situations with violation of human rights. However, the question arises of achieving a balance between moral and ethical and legal norms that exist in society in a situation where the role of these norms is decisive not only in the behavior of people, but also in the behavior of robots with artificial intelligence.

Key words: ethical norms, robots, artificial intelligence, EU legislation.

Література:

1. Mikhail, John, Universal Moral Grammar: Theory, Evidence, and the Future. Trends in Cognitive Sciences, 2007. Vol.11 No.4, Georgetown Public Law Research Paper, p. 143-152
2. Bryson, Joanna. Patience is not a virtue: the design of intelligent systems and systems of ethics. Ethics and Information Technology. (2018) 20: p. 15–26
3. Бежевець А.М. Правовий статус роботів: проблеми та перспективи визначення. Інформація і право. 2019. с. 61-67
4. Позова Д. Д. Перспективи правового регулювання штучного інтелекту за законодавством ЄС. Часопис цивілістики. 2017. № 27. 116-120.
5. Баранов О. А. Цифрова трансформація як джерело реновації парадигми рекодифікації та оновлення законодавства: огляд проблем. Право України. 2022. № 8. С. 28-47.
6. Dignum, Virginia. Ethics and Information Technology, ISSN 1388-1957, E-ISSN 1572-8439, Vol. 20, no 1, p. 1-3
7. High-Level Expert Group on Artificial Intelligence (2019) Ethics guidelines for trustworthy AI. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (дата звернення:20.10.2022)

References:

1. Mikhail, John, Universal Moral Grammar: Theory, Evidence, and the Future. Trends in Cognitive Sciences, Vol.11 No.4, Georgetown Public Law Research Paper, p. 143-152
2. Bryson, Joanna. Patience is not a virtue: the design of intelligent systems and systems of ethics. Ethics and Information Technology. (2018) 20: p. 15–26
3. Bezhevets A.M. Pravovyy status roboty: problemy ta perspektyvy vyznachennya. Informatsiya i pravo. [Legal status of robots: problems and prospects of definition. Information and law]. 2019. pp. 61-67 [in Ukrainian]
4. Pozova D. D. Perspektivy pravovoho rehulyuvannya shtuchnoho intelektu za zakonodavstvom EU. Chasopys tsyvilistyky. [Prospects of legal regulation of artificial intelligence under EU legislation. Journal of Civil Studies]. 2017. No. 27. 116-120 [in Ukrainian]

5. Baranov O. A. Digital transformation as a source of renovation of the paradigm of recodification and renewal of legislation: an overview of problems. Law of Ukraine. 2022. No. 8. P. 28-47.6. Dignum, Virginia. Ethics and Information Technology, ISSN 1388-1957, E-ISSN 1572-8439, Vol. 20, no 1, p. 1-3 [in Ukrainian]
7. High-Level Expert Group on Artificial Intelligence (2019) Ethics guidelines for trustworthy AI. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (дата звернення: 20.10.2022).

Стаття надійшла до редакції 02.12.2022

Стаття рекомендована до друку 07.12.2022