Філософсько-правові проблеми обмеження використання штучного інтелекту

Криволапов Богдан Михайлович[1]

Опубліковано	Секція	УДК
30.11.2024	Право	34.096(045)

DOI: https://doi.org/10.5281/zenodo.20455931

Анотація. *У статті досліджено філософсько-правові проблеми обмеження використання штучного інтелекту в умовах його стрімкого впровадження у різні сфери життя суспільства. Обґрунтовано, що штучний інтелект не може розглядатися лише як інструмент технічної оптимізації, оскільки його застосування безпосередньо впливає на реалізацію права на приватність, рівність, недискримінацію, доступ до правосуддя, захист персональних даних, справедливий судовий розгляд і людську гідність. Метою статті є філософсько-правовий аналіз проблеми обмеження використання штучного інтелекту на основі визначення основних ризиків його застосування для прав людини, правосуддя та публічного управління, а також обґрунтування доцільності впровадження ризик-орієнтованої моделі правового регулювання, яка заснована на принципах людської гідності, автономії, пропорційності, прозорості, підзвітності, недискримінації та ефективного контролю людини. Методологічну основу дослідження становлять діалектичний, системний, аксіологічний, формально-юридичний, порівняльно-правовий методи та метод правового прогнозування. Визначено, що ключовими ризиками застосування ШІ є алгоритмічна дискримінація, непрозорість автоматизованих рішень, порушення приватності, розмивання юридичної відповідальності та підміна людського рішення алгоритмічним прогнозом. Доведено, що правове обмеження ШІ не має зводитися до загальної заборони технології, а повинно ґрунтуватися на диференціації рівнів ризику: забороні найнебезпечніших практик, суворому контролі високоризикових систем і м'якшому регулюванні низькоризикових застосувань. Особливу увагу приділено потребі збереження реального людського контролю у правосудді та публічному управлінні. Зроблено висновок, що для України доцільним є поступове формування законодавства про ШІ з урахуванням EU AI Act, стандартів Ради Європи, рекомендацій UNESCO та національних конституційних гарантій.*

Ключові слова: *штучний інтелект, міжнародне право, міжнародно-правове регулювання, регулювання ШІ, етичні виклики, глобальна стандартизація.*

[1] кандидат юридичних наук, доцент,
доцент кафедри міжнародного приватного права,
Навчально-науковий інститут міжнародних відносин
Київського національного університету
імені Тараса Шевченка,
04119, Україна, м. Київ, вул. Юрія Іллєнка, 36/1,
kryvolapov@knu.ua,
ORCID: https://orcid.org/0000-0002-5933-7850

Philosophical and Legal Problems of Restricting the Use of Artificial Intelligence

Abstract. *The article examines the philosophical and legal problems of restricting the use of artificial intelligence in the context of its rapid implementation in various spheres of social life. It is substantiated that artificial intelligence cannot be regarded merely as a tool of technical optimization, since its application directly affects the realization of the rights to privacy, equality, non-discrimination, access to justice, personal data protection, a fair trial, and human dignity. The purpose of the article is to provide a philosophical and legal analysis of the problem of restricting the use of artificial intelligence on the basis of identifying the main risks of its application for human rights, justice, and public administration, as well as substantiating the expediency of introducing a risk-oriented model of legal regulation based on the principles of human dignity, autonomy, proportionality, transparency, accountability, non-discrimination, and effective human control. The methodological basis of the study consists of dialectical, systemic, axiological, formal-legal, comparative-legal methods, as well as the method of legal forecasting. It is determined that the key risks of applying AI include algorithmic discrimination, the opacity of automated decisions, violations of privacy, the dilution of legal liability, and the substitution of human decision-making by algorithmic prediction. It is proved that the legal limitation of AI should not be reduced to a general prohibition of the technology, but should be based on the differentiation of risk levels: the prohibition of the most dangerous practices, strict control over high-risk systems, and softer regulation of low-risk applications. Particular attention is paid to the need to preserve real human control in justice and public administration. It is concluded that, for Ukraine, the gradual formation of legislation on AI is appropriate, taking into account the EU AI Act, Council of Europe standards, UNESCO recommendations, and national constitutional guarantees.*

Keywords: *artificial intelligence, international law, international legal regulation, AI regulation, ethical challenges, global standardization.*

Постановка проблеми. Стрімке поширення систем штучного інтелекту (ШІ) у сферах державного управління, правосуддя, медицини, освіти, безпеки, фінансових послуг, інформаційних комунікацій та ринку праці актуалізує питання про межі допустимого використання таких технологій. Штучний інтелект уже не може розглядатися лише як інструмент технічної оптимізації, оскільки його застосування безпосередньо впливає на реалізацію основоположних прав людини, зокрема, права на приватність, рівність, недискримінацію, свободу, доступ до правосуддя, захист персональних даних, справедливий судовий розгляд і людську гідність [14, с. 827-828].

Відповідно, нині проблема обмеження використання штучного інтелекту має подвійний характер. З одного боку, надмірне або передчасне обмеження обсягів його застосування може стримувати інновації, знижувати конкурентоспроможність держави, ускладнювати цифрову трансформацію та перешкоджати застосуванню технологій у суспільно корисних сферах. З іншого боку, відсутність чітких меж використання ШІ створює ризики алгоритмічної дискримінації, непрозорого автоматизованого ухвалення рішень, порушення приватності, маніпуляції поведінкою людини, послаблення контролю та розмивання юридичної відповідальності [11; 12].

У філософсько-правовому вимірі, відтак, маємо говорити не лише про те, чи може держава забороняти певні технології, а, насамперед, про те, якими мають бути критерії правових обмежень використання ШІ. Такі критерії мають ґрунтуватися на принципах верховенства права, пропорційності, правової визначеності, підзвітності, недискримінації, особистісного контролю та пріоритету прав людини. Тому обмеження ШІ не повинне зводитися до загальної заборони технології як такої: його належить розуміти як формування спеціального ризик-орієнтованого правового режиму, за

якого найнебезпечніші практики забороняються, застосування високоризикових систем допускається лише за умови суворого контролю, а використання низькоризикових технологій регулюється переважно через стандарти прозорості, етичні кодекси та механізми саморегулювання [19, с. 2-5].

Аналіз останніх досліджень і публікацій. Із наукової точки зору, проблематика використання штучного інтелекту розглядається у рамках філософії, теорії права, інформаційного права, адміністративного права, цивільного права, кримінального права, міжнародного права та права Європейського Союзу. В українській науковій літературі вагоме значення мають праці О. Баранова, який аналізує поняття штучного інтелекту та витоки проблеми його правового регулювання [1]. Питання регулювання ШІ у контексті розвитку законодавства ЄС розглядають В. Пилипчук, О. Баранов та О. Гиляка, акцентуючи увагу на зв'язку між технологічною безпекою, правами людини та потребою нормативного впорядкування нових цифрових відносин [5]. О. Кожухар досліджує правове регулювання систем штучного інтелекту в ЄС [3], а О. Таран і В. Гавловський аналізують підходи ЄС та України до правового регулювання ШІ крізь призму прав людини [8]. У контексті нашого дослідження важливе значення має аналітичний матеріал «Біла книга з регулювання ШІ в Україні: бачення Мінцифри», в якому запропоновано поступовий *bottom-up* підхід до формування законодавства, орієнтований на підготовку бізнесу, розвиток саморегулювання та подальшу гармонізацію національних нормативно-правових актів з *EU AI Act* [2].

У зарубіжних наукових працях філософське-правове осмислення штучного інтелекту має місце в роботах Дж. Фон Неймана та Н. Вінера, які розглядають обчислювальні машини, кібернетику, автоматизацію та інформаційні процеси як принципово нові форми організації знання й управління [21; 22]. П. Гакер, А. Енгель і М. Мауер досліджують особливості регулювання великих генеративних моделей, наголошуючи на проблемах прозорості, відповідальності, захисту даних і розподілу обов'язків між розробниками, постачальниками та користувачами [15]. У сучасній етико-правовій літературі значущими є праці Л. Флоріді та Дж. Коулза, які обґрунтовують принципи благодіяння, нешкідливості, автономії s справедливості як основу етики ШІ [12]. Б. Міттельштадт, П. Алло, М. Таддео, С. Вахтер і Л. Флоріді аналізують етичні ризики алгоритмів, пов'язані з непрозорістю, упередженістю, відповідальністю та делегуванням рішень машинам [19].

Суттєву роль у контесті розгляду філософсько-правових проблем обмеження використання ШІ мають міжнародні та наднаціональні документи. Рекомендації UNESCO з етики штучного інтелекту закріплює людиноцентричний підхід, відповідно до якого технологічний розвиток має бути сумісним із правами людини, гідністю, недискримінацією, прозорістю та особистісним наглядом [17]. Регламент ЄС 2024/1689 щодо використання ШІ запроваджує комплексну ризик-орієнтовану модель регулювання, що включає заборонені практики, високоризикові системи, прозорість, нагляд, оцінку відповідності та спеціальні вимоги до моделей загального призначення [18]. Рамкова конвенція Ради Європи про штучний інтелект, права людини, демократію та верховенство права закріплює міжнародно-правовий підхід, у межах якого весь життєвий цикл систем ШІ має бути сумісним із правами людини, демократією та верховенством права [10].

Виділення невирішених раніше частин загальної проблеми. Попри значну кількість досліджень, що стосуються обмежень застосування ШІ на сучасному етапі, низка питань залишається недостатньо розробленою. Так, у науковій літературі ще не сформовано єдиного підходу до співвідношення етичних і правових обмежень використання ШІ; недостатньо дослідженим залишається питання меж людського контролю над автоматизованими системами; відкритою є проблема юридичної відповідальності за шкоду, завдану ШІ; потребує подальшого аналізу питання

застосування ШІ у правосудді та публічному управлінні; а конкретно для України невирішеною залишається проблема гармонізації національного законодавства про ШІ з правом ЄС і стандартами Ради Європи. Відтак, обрана проблематика потребує подальшого детального висвітлення й розгляду.

Метою статті є філософсько-правовий аналіз проблеми обмеження використання штучного інтелекту на основі визначення основних ризиків його застосування для прав людини, правосуддя та публічного управління, а також обґрунтування доцільності впровадження ризик-орієнтованої моделі правового регулювання, яка заснована на принципах людської гідності, автономії, пропорційності, прозорості, підзвітності, недискримінації та ефективного контролю людини.

Методологічну основу дослідження становить сукупність загальнонаукових, філософських і спеціально-правових методів. Діалектичний метод використано для аналізу суперечностей між інноваційним потенціалом ШІ та ризиками його неконтрольованого застосування. Системний метод дав змогу розглянути ШІ як елемент ширшої соціально-правової системи, що включає технології, дані, правові норми, інститути контролю, суб'єктів відповідальності та користувачів. Аксіологічний метод застосовано для оцінки ШІ крізь призму людської гідності, свободи, рівності та справедливості. Формально-юридичний метод використано для аналізу законодавства України, Регламенту ЄС 2024/1689, Рекомендацій UNESCO та Рамкової конвенції Ради Європи. Порівняльно-правовий метод дав змогу зіставити український підхід до регулювання ШІ з європейською ризик-орієнтованою моделлю. Метод правового прогнозування використано для визначення можливих напрямів розвитку національного законодавства у сфері ШІ в майбутньому та його гармонізації з правом ЄС.

Виклад основного матеріалу дослідження. Штучний інтелект є однією з ключових технологій сучасного цифрового суспільства, однак його правова природа залишається складною та багатовимірною [1, с. 32-33]. Згідно з підходом UNESCO, системи ШІ можуть обробляти дані та інформацію у спосіб, що нагадує інтелектуальну поведінку, включаючи елементи навчання, міркування, прогнозування, сприйняття, планування або контролю [13, с. 25]. Тобто, ШІ не зводиться лише до машинного навчання, а також охоплює знання, логічне міркування, планування, оптимізацію, пошук, обробку природної мови та інші технологічні підходи.

Філософське осмислення ШІ пов'язане з питанням визначення межі між людським і машинним інтелектом. На ранніх етапах розвитку обчислювальної техніки комп'ютери розглядалися насамперед як машини для виконання складних обчислень і зберігання великих масивів інформації. Уже на той час стало очевидним, що в окремих сферах – обчисленні, пам'яті, швидкості обробки даних – машина може істотно перевищувати людські можливості. Саме цей факт породив припущення про можливість створення штучного інтелекту, який за своїми властивостями наблизиться до людського розуму. Так, Дж. Фон Нейман порівнював принципи роботи обчислювальних машин і мозку, наголошуючи на значенні формальних, логічних та обчислювальних процесів [21]. А Н. Вінер, розвиваючи кібернетику, показав, що управління, зворотний зв'язок і комунікація є спільними для живих організмів, машин і соціальних систем [22].

Хоча, якщо машина здатна обробляти інформацію, робити прогнози й адаптуватися до даних, це не означає, що вона має свідомість, моральну автономію або відповідальність. Саме ця відмінність має принципове правове значення. Людина є носієм гідності, свободи, прав і обов'язків, тоді як ШІ залишається технологічною системою, створеною, навченою, впровадженою і застосованою людьми. Тому філософсько-правова проблема полягає не в тому, щоб визнати ШІ самостійним

суб'єктом права, а в тому, щоб визначити, які індивіди та організації відповідають за його створення, використання, помилки й наслідки.

Значна частина суспільних побоювань щодо ШІ пов'язана з ідеєю можливого виходу штучного інтелекту з-під контролю людини. В. Віндж, аналізуючи перспективу технологічної сингулярності, припускав, що створення інтелекту, який перевищить людський розум, може радикально змінити майбутнє людства [20]. Проте в сучасному правовому аналізі більшу практичну небезпеку становить не абстрактна поява «свідомої машини», а конкретні ризики вже наявних систем: непрозоре автоматизоване оцінювання, дискримінаційні алгоритми, маніпулятивні інтерфейси, масове спостереження, неконтрольована обробка персональних даних, помилкові прогнози у правосудді та делегування публічно-владних рішень технічним системам.

У цьому контексті особливого значення набуває категорія людської гідності. Ст. 3 Конституції України визнає людину, її життя і здоров'я, честь і гідність, недоторканність і безпеку найвищою соціальною цінністю [4]. Відтак, будь-яке використання ШІ має бути підпорядковане не лише цілям ефективності, економії ресурсів або автоматизації, а й вимозі поваги до людини як автономного суб'єкта. Якщо ШІ використовується для прийняття рішень щодо доступу до роботи, освіти, соціальної допомоги, медичних послуг або судового захисту, то людина не може бути зведена до набору даних, статистичного профілю чи алгоритмічної категорії.

Правовий аналіз обмеження ШІ доцільно здійснювати через призму кількох ключових ризиків. Першим є ризик порушення приватності та захисту персональних даних. Системи ШІ часто потребують великих масивів даних, серед яких можуть траплятися чутливі персональні дані, біометрична інформація, дані про поведінку, пересування, соціальні зв'язки, політичні погляди, стан здоров'я або економічний статус. Закон України «Про захист персональних даних» від 1 червня 2010 року закріплює вимоги до обробки персональних даних, зокрема законність, визначеність мети, пропорційність і захист від незаконного доступу [7]. Однак ШІ ускладнює застосування традиційних принципів захисту даних, оскільки система може виявляти нові кореляції, виводити приховані характеристики особи та створювати профілі, які сама людина прямо не повідомляла.

Другим ризиком є алгоритмічна дискримінація. Алгоритми машинного навчання залежать від даних, на яких вони навчаються. Якщо такі дані відображають історичні нерівності, соціальні упередження або неповноту інформації, ШІ може відтворювати й посилювати дискримінаційні практики. Наприклад, система оцінювання кандидатів на роботу може віддавати перевагу певним групам, якщо попередні дані про успішних працівників були сформовані у дискримінаційному середовищі. Система оцінки кредитоспроможності може непрямо дискримінувати осіб за місцем проживання, віком або соціальним походженням. Тому правове обмеження ШІ має включати не лише формальну заборону дискримінації, а й обов'язок перевірки наборів даних, тестування моделей, документування ризиків і створення процедур оскарження алгоритмічних рішень [11, с. 90-91; 19].

Третім ризиком є непрозорість і неможливість роз'яснення алгоритмічних рішень. Багато сучасних систем машинного навчання функціонують як «чорні скриньки», тобто дають результат без зрозумілого пояснення логіки його формування. У технічних сферах така непрозорість може бути прийнятною, якщо отриманий результат легко перевірити. Однак, у правосудді, медицині, публічному управлінні чи кредитуванні неможливість пояснити наявні рішення може порушувати право особи знати підстави втручання у її права. Як результат, функціонування ШІ, яке можна пояснити, є не лише технічним, а й правовим принципом: особа повинна мати можливість зрозуміти, чому щодо неї ухвалено певне рішення, які дані були використані, які критерії застосовано та як це рішення можна оскаржити [11].

Четвертим ризиком є розмивання відповідальності. Коли певне рішення формується складною системою, у створенні якої беруть участь розробники, постачальники даних, інтегратори, користувачі, оператори та органи влади, виникає небезпека «розсіювання» відповідальності. Кожен учасник може стверджувати, що не контролював остаточний результат. У зв'язку з цим, правове регулювання має визначати не абстрактну «відповідальність ШІ», а конкретні обов'язки суб'єктів життєвого циклу системи – розробника, постачальника, імпортера, розповсюджувача, користувача, органу публічної влади та особи, яка ухвалює остаточне рішення [18].

П'ятим ризиком є підміна людського рішення алгоритмічним прогнозом. ШІ може бути корисним інструментом підтримання процесу прийняття рішень, однак доволі небезпечним є перетворення його рекомендації на фактично обов'язковий висновок. Наприклад, даний нюанс є особливо важливим у судовій сфері, оскільки суддя має оцінювати докази, застосовувати право і мотивувати рішення як незалежний суб'єкт правосуддя. Алгоритм може допомагати в пошуку практики, систематизації доказів або виявленні типових справ, однак не може замінити судове переконання, принцип змагальності, право бути почутим і обов'язок мотивувати рішення.

Міжнародна практика вже містить приклади використання ШІ у правовій сфері. Юристи використовують ШІ для правового пошуку, аналізу судової практики, прогнозування результатів справ і підготовки процесуальних документів. Так, в Аргентині у певних категоріях справ застосовувалися системи автоматизації підготовки судових рішень; у Китаї Інтернет-суд м. Ханчжоу впроваджував систему аналізу доказів із використанням блокчейну, великих даних, хмарних обчислень і ШІ [13, с. 43; 23]. У Мексиці система *Expertius* використовувалася для підтримки визначення права на соціальне забезпечення на підставі попередніх справ, матеріалів слухань і рішень [9]. У Колумбії система *PretorIA* розроблялася для допомоги Конституційному суду у відборі справ, не замінюючи суддів, а оптимізуючи аналіз матеріалів [9; 13, с. 43].

Ці приклади демонструють, що використання ШІ у правосудді може мати позитивний ефект: зменшення навантаження на суддів, пришвидшення пошуку релевантної інформації, уніфікація типових процесів, покращення доступу до правової інформації. Водночас, вони виявляють і фундаментальну межу: автоматизація допоміжних операцій є допустимою лише тоді, коли остаточне рішення залишається за людиною, а учасники процесу мають можливість зрозуміти, перевірити й оскаржити використання алгоритмічних інструментів.

Відтак, особливе значення нині має концепція *human-in-the-loop*, згідно з якою ШІ може формувати рекомендацію, але остаточне рішення ухвалює компетентна людина. У сфері охорони здоров'я, наприклад, це означає, що алгоритм може допомагати лікарю з діагностикою, але відповідальність за медичне рішення залишається за лікарем. У правосудді цей алгоритм може допомагати судді систематизувати матеріали, але не може самостійно визначати міру провини, розмір покарання або результат спору. Концепція *human-on-the-loop* передбачає нагляд людини за системою ШІ з можливістю втручання, тоді як підхід *human-out-of-the-loop* означає відсутність реального людського контролю [16; 13, с. 43-45]. На нашу думку, для високоризикових сфер на зразок юриспруденції остання модель є неприйнятною або потребує особливо суворих обмежень.

Відповідно до положень Регламенту ЄС 2024/1689, на рівні даної організації закріплено ризик-орієнтований підхід, корий не забороняє ШІ як технологію, проте диференціює правовий режим його застосування залежно від рівня ризику, тобто на парктиці забороняє найнебезпечніші практики (маніпулятивні або оманливі техніки, що істотно спотворюють поведінку людини; певні види прогнозування ризику вчинення кримінального правопорушення лише на підставі профілювання; нецільове

збирання зображень облич для створення баз розпізнавання; біометрична категоризація для виведення чутливих ознак особи тощо) [18]. Високоризикові системи ШІ за Регламентом ЄС не забороняються, але допускаються лише за умови виконання спеціальних вимог, до яких належать система управління ризиками, належне управління даними, технічна документація, прозорість для користувача, індивідуальний контроль, точність, стійкість і кібербезпека [18]. Слід зауважити, що такий підхід має очевидне філософсько-правове підґрунтя, адже забороняє не використання технологій як таких, а практик, які явно суперечать людській гідності, автономії, рівності та приватності. Для України дана модель має суттєве значення, оскільки відповідає європейському інтеграційному курсу нашої держави та потребі гармонізації національного законодавства з правом ЄС. Крім того, вона може бути адаптована до українських умов через поетапне запровадження регулювання, що передбачено Білою книгою Мінцифри [2], створюючи рамку для секторального регулювання: окремі правила можуть бути потрібні для правосуддя, медицини, освіти, оборони, фінансового сектору, публічного управління та критичної інфраструктури.

Рамкова конвенція Ради Європи про ШІ збагачує і доповнює згаданий підхід, надаючи йому міжнародно-правового виміру. Її положення ґрунтуються на переконанні, що діяльність у межах життєвого циклу систем ШІ має відповідати правам людини, демократії та верховенству права, виходячи з принципів людської гідності й автономії, рівності й недискримінації, приватності й захисту персональних даних, прозорості й нагляду, відповідальності, надійності та безпечних інновацій [10].

У філософському аспекті обмеження ШІ можна обґрунтувати через принцип автономії людини. Якщо система ШІ маніпулює поведінкою особи, приховано впливає на її вибір, використовує психологічні вразливості або формує інформаційне середовище так, що людина втрачає здатність до самостійного судження, така система порушує її автономію. Відтак, заборона маніпулятивних практик є не технічним, а антропологічним обмеженням: держава повинна захищати здатність людини бути автором власних рішень. Іншим філософським критерієм є справедливість. Алгоритмічні системи часто позиціонуються як нейтральні, оскільки вони працюють із даними та формальними моделями. Однак формальність не гарантує справедливості: якщо дані неповні, історично упереджені або неправильно інтерпретовані, алгоритм може створити видимість об’єктивності там, де фактично відбувається відтворення нерівності. Тому справедливість ШІ вимагає не лише однакового застосування алгоритму до всіх, а й перевірки того, чи не призводить такий алгоритм до непропорційного негативного впливу на окремі групи осіб.

Принцип пропорційності є центральним для правового обмеження ШІ, адже будь-яке обмеження має бути придатним для досягнення легітимної мети, необхідним і співмірним. Наприклад, автоматизоване оцінювання ризику особи у кримінальному провадженні може бути корисним як допоміжний інструмент, але не може бути єдиною підставою для обмеження свободи. Проблема відповідальності за шкоду, спричинену ШІ, також потребує оновлення традиційних юридичних конструкцій. У цивільному праві відповідальність зазвичай пов’язана з наявністю шкоди, протиправності, причинного зв’язку та вини, проте у випадку ШІ причинний зв’язок може бути складним для доведення через багаторівневість технологічного процесу. До нанесення шкоди можуть призвести помилки в даних, недоліки моделі, неправильне налаштування, неналежне використання, відсутність оновлення або кібератака. Відтак, перспективним є поєднання кількох моделей відповідальності: відповідальності розробника за дефекти системи, відповідальності користувача за неналежне застосування, відповідальності органу влади за незаконне рішення, а також спеціальних режимів підвищеної відповідальності для високоризикових систем.

Окремим аспектом під час розгляду питання правового обмеження ШІ є інтелектуальна власність. Генеративні моделі здатні створювати тексти, зображення, музику, програмні коди та інші об'єкти, що викликає питання про правовий режим результатів, використання захищених творів для навчання моделей і межі їх добросовісного використання. Закон України «Про авторське право і суміжні права» виходить із антропоцентричного розуміння творчості, за якого автором є фізична особа [6]. Відповідно, результати, повністю згенеровані ШІ без творчого внеску людини, не можуть автоматично прирівнюватися до творів людини. Водночас, якщо людина здійснює творчий відбір, редагування, композицію або концептуальне керівництво, питання правової охорони має вирішуватися з урахуванням реального людського внеску.

У сфері публічного управління ШІ може бути ефективним інструментом аналізу великих масивів даних, прогнозування потреб населення, виявлення корупційних ризиків, оптимізації адміністративних послуг тощо, але застосування ШІ державою має бути відкритим, нормативно визначеним і передбачати ефективний механізм оскарження. В даному контексті слід відзначити, що особливо чутливою щодо питання застосування ШІ є сфера судочинства. Право на справедливий суд передбачає незалежність і неупередженість суду, доступ до доказів, мотивованість рішення, право бути почутим і можливість оскарження. Якщо судове рішення фактично ґрунтується на непрозорому алгоритмі, сторона не може ефективно заперечити його логіку, що підриває принцип змагальності та рівності сторін. Тому використання ШІ в судовому порядку має бути обмежене допоміжними функціями: пошуком практики, класифікацією справ, управлінням документообігом, попереднім аналізом доказів, виявленням процесуальних строків. Натомість, остаточне вирішення прав і обов'язків особи має залишатися виключно за суддею.

Отже, філософсько-правова проблема обмеження застосування ШІ полягає не в протиставленні людини й технології, а в побудові такого режиму, за якого технологія служить людині, не підміняючи її правову, моральну й політичну суб'єктність. Відповідно, для України доцільним є поетапне формування правового режиму ШІ. На першому етапі необхідно закріпити базові принципи: людиноцентричність, безпека, прозорість, недискримінація, захист персональних даних, людський контроль, підзвітність і право на оскарження автоматизованого рішення. На другому етапі слід визначити категорії ризику та сфери, у яких ШІ вважається високоризиковим. На третьому етапі потрібно встановити процедури оцінки відповідності, аудиту, реєстрації високоризикових систем, сертифікації та постійного моніторингу. На четвертому етапі варто передбачити спеціальні правила для генеративного ШІ, зокрема, щодо маркування синтетичного контенту, захисту авторських прав, протидії дезінформації та відповідальності за шкідливі результати [3; 8 ,с. 63].

У цьому контексті варто розмежовувати етичні, технічні та правові обмеження ШІ. Етичні принципи формують ціннісний горизонт – гідність, справедливість, нешкідливість, автономія, доброчесність. Технічні обмеження забезпечують реалізацію цих принципів через архітектуру системи – аудит даних, тестування, захист від атак, логування, документацію, пояснюваність, контроль якості. Правові обмеження перетворюють цінності й технічні стандарти на обов'язки, за порушення яких настає юридична відповідальність. Тобто, без правового механізму етика ризикує залишитися декларацією; без етики право може стати формальним; а без технічної реалізації обидва рівні будуть неефективними. Відтак, обмеження використання ШІ слід розуміти не як заперечення технологічного прогресу, а як правову форму його гуманізації. Технологія, що здатна впливати на свободу, приватність, рівність і доступ до правосуддя, не може залишатися поза межами права. Водночас, право не повинне бути суто репресивним або заборонним. Його завдання полягає у створенні збалансованої

моделі, яка дозволяє використовувати переваги ШІ, але запобігає перетворенню людини на об'єкт алгоритмічного управління.

Висновки. Штучний інтелект є не лише технологічним, а й філософсько-правовим феноменом, оскільки його використання безпосередньо зачіпає питання людської гідності, автономії, свободи, рівності, справедливості, приватності та відповідальності. Основна проблема полягає не в існуванні ШІ як такого, а в тих формах його застосування, які можуть призводити до дискримінації, маніпуляції, непрозорого ухвалення рішень, незаконного втручання у приватне життя та підміни людського судження машинним прогнозом.

Обмеження використання ШІ має бути ризик-орієнтованим, пропорційним і юридично визначеним. Найнебезпечніші практики, які порушують людську гідність, автономію, рівність і приватність, мають бути заборонені. Високоризикові системи повинні допускатися лише за умови дотримання вимог до управління ризиками, якості даних, технічної документації, прозорості, людського нагляду, точності, кібербезпеки й підзвітності. Низькоризикові застосування можуть регулюватися м'якшими засобами, зокрема, через стандарти, кодекси поведінки та механізми саморегулювання. У правосудді та публічному управлінні використання ШІ має бути особливо обережним. Алгоритмічні системи можуть виконувати допоміжні функції, однак не повинні замінювати суддю, адміністративний орган або іншу компетентну особу у прийнятті рішень, що істотно впливають на права людини. Формальне включення людини в процес недостатнє: людський контроль має бути реальним, компетентним і здатним змінити або скасувати результат роботи системи.

Для України доцільним є поступове формування законодавства про ШІ з урахуванням положень *EU AI Act*, Рамкової конвенції Ради Європи, Рекомендацій UNESCO та національних конституційних гарантій, що має поєднувати розвиток інновацій із захистом людини від алгоритмічної сваволі. У цьому полягає головний філософсько-правовий сенс обмеження використання штучного інтелекту: не зупинити технологічний прогрес, а підпорядкувати його праву, свободі та людській гідності.

Список використаних джерел

1. Баранов О. А. Визначення терміну «штучний інтелект». *Інформація і право*. 2023. № 1(44). С. 32-49.
2. Біла книга з регулювання ШІ в Україні: бачення Мінцифри. Київ: Міністерство цифрової трансформації України, 2024. 30 с. URL: https://storage.thedigital.gov.ua/files/a/ba/d5da75c2613e331bb89258f950adcbae.pdf (дата звернення: 28.07.2024).
3. Кожухар О. Г. Правове регулювання систем штучного інтелекту в ЄС: передумови, сучасний стан та перспективи. *Наукові записки НаУКМА. Серія «Юридичні науки»*. 2024. Т. 13. С. 65-73.
4. Конституція України: Закон України від 28.06.1996 № 254к/96-ВР. Дата оновлення: 01.01.2020. *Верховна Рада України: Офіційний вебпортал Парламенту України*. URL: https://zakon.rada.gov.ua/laws/show/254к/96-вр (дата звернення: 30.07.2024).
5. Пилипчук В. Г., Баранов О. А., Гиляка О. С. Проблема правового регулювання у сфері штучного інтелекту в контексті розвитку законодавства Європейського Союзу. *Вісник Національної академії правових наук України*. 2022. Т. 29, № 2. С. 35-62.

6. Про авторське право і суміжні права: Закон України від 01.12.2022 № 2811-IX. Дата оновлення: 15.04.2023. *Верховна Рада України: Офіційний вебпортал Парламенту України*. URL: https://zakon.rada.gov.ua/go/2811-20 (дата звернення: 30.07.2024).
7. Про захист персональних даних: Закон України від 01.06.2010 № 2297-VI. Дата оновлення: 22.02.2024. *Верховна Рада України: Офіційний вебпортал Парламенту України*. URL: https://zakon.rada.gov.ua/laws/show/2297-17#Text (дата звернення: 30.07.2024).
8. Таран О. В., Гавловський В. Д. Правове регулювання штучного інтелекту в Європейському Союзі та Україні: основні підходи та права людини. *Інформація і право*. 2024. № 1(48). С. 62-67.
9. Conoce nuestra investigación sobre PretorIA, la tecnología que incorpora la Inteligencia Artificial a la Corte Constitucional. *Dejusticia*. 19.04.2021. URL: https://www.dejusticia.org/conoce-nuestra-investigacion-sobre-pretoria-la-tecnologia-que-incorpora-la-inteligencia-artificial-a-la-corte-constitucional/ (дата звернення: 28.08.2024).
10. Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law of 5 September 2024. *Council of Europe*. URL: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence (дата звернення: 28.09.2024).
11. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI / A. Barredo Arrieta et al. *Information Fusion*. 2020. Vol. 58. P. 82-115.
12. Floridi L., Cowls J. A Unified framework of five principles for AI in society. *Harvard Data Science Review*. 02.07.2019. URL: https://hdsr.mitpress.mit.edu/pub/l0jsh9d1/release/8 (дата звернення: 29.08.2024).
13. Global toolkit on AI and the rule of law for the judiciary / UNESCO. Paris: UNESCO, 2023. 212 p. URL: https://unesdoc.unesco.org/ark:/48223/pf0000387331 (дата звернення: 28.08.2024).
14. Goretty C., Martinez B. La inteligencia artificial y su aplicación al campo del Derecho. *Alegatos*. 2012. No. 82. P. 827-846.
15. Hacker P., Engel A., Mauer M. Regulating ChatGPT and other large generative AI models. *Cornell University*. 12.06.2023. URL: https://arxiv.org/abs/2302.02337 (дата звернення: 29.08.2024).
16. Model Artificial Intelligence Governance Framework. 2nd ed. Singapore: Personal Data Protection Commission, Infocomm Media Development Authority, 2020. 68 p. URL: https://www.imda.gov.sg/-/media/imda/files/infocomm-media-landscape/sg-digital/tech-pillars/artificial-intelligence/second-edition-of-the-model-ai-governance-framework.pdf (дата звернення: 18.09.2024).
17. Recommendation on the Ethics of Artificial Intelligence of 23 November 2021 / UNESCO. Paris: UNESCO, 2022. 43 p. URL: https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence (дата звернення: 28.08.2024).
18. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence. *Official Journal of the European Union*. 2024. L Series. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (дата звернення: 28.08.2024).
19. The ethics of algorithms: Mapping the debate / B. D. Mittelstadt et al. *Big Data & Society*. 2016. Vol. 3, Issue 2. P. 1-21.
20. Vinge V. The coming technological singularity: How to survive in the post-human era. *Whole Earth Review*. 1993. No. 2. URL: https://users.manchester.edu/Facstaff/SSNaragon/Online/100-FYS-

F15/Readings/Vinge,%20The%20Coming%20Technological%20Singularity.pdf (дата звернення: 20.09.2024).
21. Von Neumann J. The computer and the brain. New Haven: Yale University Press, 1958. 82 p.
22. Wiener N. Cybernetics: Or control and communication in the animal and the machine. Cambridge: MIT Press, 1948. 194 p.
23. Xuan H. One-click access to evidence analysis results. Hangzhou Internet Court launches intelligent evidence analysis system. *China Courts Network*. 17.08.2019. URL: https://www.chinacourt.org/article/detail/2019/12/id/4747683.shtml (дата звернення: 30.09.2024).