# РОЗДІЛ XXI. ЕТИЧНІ АСПЕКТИ ВИКОРИСТАННЯ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПРОТИДІЇ КОРУПЦІЇ В ЕКОНОМІЦІ ТА ЗАХИСТУ ПРАВ ЛЮДИНИ

DOI 10.33407/lib.NAES.id/750378

**Заярний Олег**[1[0000-0003-4549-7201]], **Остапенко Роман**[1[0009-0005-1996-3450]],
**Почаєвець Олена**[1[0000-0003-4521-9922]], **Чуян Андрій**[1[0009-0005-4089-4795]]
[1]Київський національний університет імені Тараса Шевченка
*olehzaiarnyi@knu.ua, roman.ostapenko@knu.ua, british_idler@knu.ua, olena.pochaievets@knu.ua*

**Анотація**. Цей розділ монографії, присвячений аналізу етичних аспектів використання технологій штучного інтелекту для протидії корупції в економіці та забезпечення захисту прав людини, піднімає питання людиноорієнтованого використання технологій штучного інтелекту в питаннях боротьби з корупцією. Також, для всебічного розкриття проблематики, у розділі розкрито деякі технічні аспекти використання моделей генеративного штучного інтелекту, наведено конкретні приклади, на основі яких проаналізовано чинні тенденції та потенційні можливості розвитку цієї технології. Усе це вказується як для європейського, так і для українського економічного простору. Водночас, співавтори цієї роботи дотримуються наукової позиції, відповідно до якої, з дедалі більшою інтеграцією України в європейське правове поле наведені перспективи застосування штучного інтелекту в цілях протидії корупції зберігають свою актуальність для всіх держав-членів Ради Європи. Цю частину монографії написано у рамках НДР «Розробка інструментів для розвитку смарт-інфраструктури розумних міст з урахуванням викликів війни та повоєнного відновлення України» Д/Б теми № 25БП013-01М, МОН України, номер держ. реєстрації 0125U001221).

**Ключові слова:** генеративні моделі штучного інтелекту; економічна безпека; запобігання корупції; права людини; публічна смарт-інфраструктура; штучний інтелект.

# ETHICAL ASPECTS OF USING ARTIFICIAL INTELLIGENCE TECHNOLOGIES FOR COUNTERING CORRUPTION IN THE ECONOMY AND PROTECTING HUMAN RIGHTS

**Zaiarnyi Oleh, Ostapenko Roman, Pochaievets Olena, Chuian Andrii**

Taras Shevchenko National University of Kyiv

**Summary:** This monograph chapter, dedicated to analyzing the ethical aspects of using artificial intelligence technologies to counter corruption in the economy and ensure the protection of human rights, addresses the human-centric application of AI technologies in anti-corruption efforts. Furthermore, for a comprehensive exploration of the subject, the chapter examines certain technical aspects of utilizing generative artificial intelligence models, provides specific examples, and analyzes current trends and the potential for this technology's development based on these examples. This analysis covers both the European and Ukrainian economic spaces. Concurrently, the co-authors maintain the academic position that, given Ukraine's increasing integration into the European legal framework, the outlined prospects for applying artificial intelligence in anti-corruption efforts remain relevant for all member states of the Council of Europe. This part of the monograph was written within the framework of the research project "Development of Tools for Smart City Infrastructure Advancement Considering the Challenges of War and Post-War Recovery in Ukraine" (State Budget theme No. 25БП013-01М, Ministry of Education and Science of Ukraine, State Registration No. 0125U001221).

**Key words:** generative models of artificial intelligence; economic security; corruption prevention; human rights; public smart infrastructure; artificial intelligence.

**Вступ**. Аналогічно до багатьох держав світу, в Україні корупція залишається однією з головних перешкод на шляху сталого розвитку, підриваючи довіру до державних інститутів, спотворюючи ринкові механізми та порушуючи права людини.

Згідно з Індексом сприйняття корупції Transparency International, багато країн світу, включаючи Україну, продовжують стикатися з високим рівнем корупції [1].

Використання технологій штучного інтелекту (ШІ) відкриває нові можливості для боротьби з корупцією, дозволяючи автоматизувати аналіз великих обсягів даних, виявляти аномалії та прогнозувати потенційні корупційні ризики.

Водночас, застосування ШІ також породжує низку етичних та правових питань, пов'язаних з прозорістю алгоритмів, захистом персональних даних та можливими дискримінаційними результатами виконання поставлених завдань. Такі ризики певною мірою можуть стосуватися вразливості ШІ до кібератак і витоку конфіденційних даних, неточної оцінки корупціогенних факторів у конкретному юридичному контексті, тощо.

Міжнародні акти, такі як Конвенція ООН проти корупції [2], закликають держави-учасниці до використання сучасних технологій для запобігання та боротьби з корупцією.

Проте, для ефективного застосування ШІ у протидії корупційним правопорушенням необхідно враховувати принципи захисту прав людини, закріплених в таких документах, як Загальна декларація прав людини [3], а також Конвенція про захист прав людини і основоположних свобод [4].

На національному рівні, Законом України "Про запобігання корупції" [5] передбачено використання інформаційних технологій для виявлення та запобігання корупційним правопорушенням. Поряд з цим, згаданий Закон України, так само, як і інші акти національного законодавства у сфері протидії корупції не передбачають підстав, умов і процедур застосування ШІ для відповідних цілей.

У зв’язку з цим, поза правовим регулюванням залишаються правові засади застосування ШІ для протидії корупції, правові і технічні вимоги до моделей ШІ, які можуть потенційно використовуватися з метою виявлення корупційних ризиків, вимоги до завдань, які можуть виконуватися з використанням цифрових рішень на основі ШІ, а також спеціальний механізм захисту прав людини у відповідній сфері правового регулювання.

**Аналіз літератури та постановка проблеми**

Так, Дітер Ціннбауер у своїй праці аналізує сучасний стан застосування ШІ у боротьбі з корупцією, висвітлюючи як досягнення, так і виклики у цій сфері. На думку дослідника, ШІ продемонстрував ефективність у таких областях, як забезпечення прозорості державних закупівель, дотримання нормативних вимог, виявлення шахрайства та боротьба з відмиванням грошей. Зокрема, ШІ сприяє ідентифікації нових індикаторів ризику, аналізу великих обсягів даних та виявленню складних схем змови. Водночас, автор вважає, що незважаючи на досягнення, існують перешкоди, такі як обмежені ресурси, низька якість даних, організаційний опір та цифровий розрив. Також ШІ-системи можуть мати вроджені упередження, генерувати помилкові результати та страждати від недостатньої прозорості алгоритмів [6].

У іншій публікації, так само зроблений значний акцент на перевагах і ризиках застосування ШІ для протидії корупції в економіці. Проте, в роботі детально аналізуються окремі групи ризиків, пов’язаних з інтеграцією ШІ у антикорупційні стратегії різних держав. Авторами наголошується на тому, що ефективність ШІ залежить від якості, повноти та надійності даних. Неповні або упереджені дані можуть призвести до неточних результатів та підриву довіри до антикорупційних заходів. Поряд з цим, збір та аналіз великих обсягів даних можуть порушувати права на приватність, викликаючи етичні питання щодо використання особистої інформації [7].

У роботі, присвяченій застосуванню ШІ, зокрема, ChatGPT у навчанні майбутніх правників, О. А. Заярний та С. І. Ілларіонова звертають окрему увагу на необхідність опанування студентами правниками додаткових знань щодо можливості застосування генеративних моделей ШІ для запобігання корупційним правопорушенням. Поряд з іншими напрямами підготовки юристів з використанням ШІ, вчені-правники формулюють переваги і недоліки застосування ШІ для виконання різних завдань на публічній службі, формулюють приклади інструкцій з оформлення запитів на пошук інформації, яка може вказувати на ознаки корупційних ризиків, правопорушень, тощо [8, с. 165-168].

Метою цієї роботи є дослідження етичних аспектів використання технологій ШІ для протидії корупції в економіці та захисту прав людини. У цій роботі буде проаналізовано можливості та ризики застосування ШІ в антикорупційній діяльності, розглянуто зарубіжний та національний досвід, а також запропоновано рекомендації щодо забезпечення етичного та правового регулювання використання ШІ у цій сфері.

**Результати дослідження.** На момент публікації, а саме 2025 рік, у серед науковців відсутнє універсальне визначення ШІ.

В межах цієї роботи ми будемо виходити з того, що ШІ – це система на основі машинного навчання, що робить висновки та генерує рішення на основі вхідних даних [6]. Також, попри активне використання цього терміну, ми акцентуємо увагу на тому, що наразі мова йде про так званий слабкий штучний інтелект або вузького призначення (Artificial Narrow Intelligence, ANI) [10] готового хоч і для багатьох, але по факту специфічних або вузькопрофільних завдань. Натомість сильний, або загальний, штучний інтелект (Artificial General Intelligence, AGI) [10] – це те, що може з’явитися в майбутньому, враховуючи сьогоденні темпи розвитку технології, однак досі залишається нереалізованою розробкою. Одним із критеріїв сильного штучного інтелекту може бути самосвідомість, що підніматиме ряд інших питань у етичному та моральному планах (особливо в контексті протидії корупції). Однак у подальшому тексті розглянуто саме ШІ прикладного типу.

На сьогоднішній день загально доступні генеративні моделі ШІ можуть бути використані для виявлення корупційних схем, аналізу даних про державні закупівлі, податкові декларації, фінансові операції та інші сфери економічної діяльності.

Застосування ШІ дозволяє автоматизувати процеси виявлення корупції, підвищити їх швидкість та точність, а також виявляти складні корупційні схеми, які важко виявити традиційними методами.

У світі вже є приклади успішного застосування ШІ для протидії корупції. Серед прикладів успішного застосування цієї технології в механізмах протидії корупції можна, зокрема, вказати на наступні:

**Естонія:** ШІ використовується для аналізу державних закупівель. Також важливо зазначити досвід додатку з інтегрованими ШІ-можливостями [11], а саме доступ до державних сервісів через Bürokratt [12] (своєрідний аналог української "Дії" [13]) – усе це, з використанням сервісів анонімізації, аналізу даних та відкритих реєстрів даних має слугувати якісним підґрунтям для боротьби з корупцією. Не варто забувати і про взаємодію з такими техногігантами у царині ШІ, як OpenAI – навіть ті апробації AI-технологій, які здійснюватимуться в школах [14] можуть розвинути довіру в населення до цих технологій, а також показати всі складнощі, з якими може зіткнутися уряд при подальшому провадженню, вже конкретно за напрямом протидії корупції.

**Данія:** ШІ застосовується для аналізу податкових декларацій [15] та захисту цих даних [16]. Загалом подібний підхід стає розповсюдженим серед європейських країн. Конкретно Данія таким чином відслідковує більше половини випадків податкового шахрайства, що, очевидно, дозволяє оподаткувати мільярди євро коштів, котрі могли безкарно піти в тіньовий бізнес [17].

**Німеччина:** нещодавно займалась дослідженням, як ШІ може використовуватися для фінансових розслідувань [18]. Подібно до інших країн ЄС, тут активний розвиток стратегії та впровадження ШІ у державні сфери життя розпочався близько 2018 року [19].

**Франція:** ШІ застосовується для моніторингу державних контрактів [20]. Необхідно зазначити, що також французький уряд займається розробкою стратегії використання ШІ в галузі "розумних міст" (smart cities) [21]. Таким чином, подібні моделі можуть використовуватися не лише на рівні опрацювання документації, а й загалом протидіяти корупційній діяльності на низовому рівні (шляхом аналізу деперсоніфікованої інформації одержаної в межах міст, районів та навіть вулиць).

**Сінгапур:** Використання ШІ для аналізу державних закупівель та виявлення потенційних конфліктів інтересів. Загалом Сінгапур активно впроваджує ШІ у сферу надання державних послуг та надає перевагу практичності перед зарегульованістю нової технології [22]. Одним із таких прикладів є AISG [23] - національна програма для об'єднання інституцій у розробці та впровадженні ШІ-технологій у побут громадян (для державного адміністрування, навчання, створення AI-продуктів тощо) [24].

**Бразилія:** Застосування ШІ для моніторингу державних видатків та виявлення шахрайства. До прикладу, Національна конфедерація страхових компаній користується ШІ-аналітикою для запобігання шахрайству [25]. Цікавий також приклад аналізу імпорту та експорту з використанням HARPIA - аналізатора ризиків та прикладного штучного інтелекту [26]. Однак, подібні технології розвиваються і в негативному контексті - подекуди ШІ використовують саме для шахрайства і боротьба з цим триває [27].

**США:** Використання ШІ для аналізу фінансових транзакцій та виявлення випадків відмивання грошей [28]. Загалом Сполучені Штати є наразі чи не головним рушієм прогресу в цій царині, як один із учасників "холодної війни за штучний інтелект" (більш коректно "ШІ-холодна війна" - AI Cold War). Основними гравцями у приватному секторі тут є OpenAI (ChatGPT [29], Whisper для транскрибування аудіо [30]), Google (Gemini [31], Gemma [32], створення архітектури Transformer [33]), Microsoft (Copilot [34], Phi [35]), Meta (Llama [36]). Обов'язково треба згадати в цьому дискурсі Stargate Project із обсягом фінансування 500 мільярдів доларів США - амбітну мету з прискорення розробки ШІ, до якої залучено головних техногігантів галузі [37].

В Україні застосування ШІ для протидії корупції перебуває на початковій стадії. Існують окремі ініціативи, спрямовані на використання ШІ для аналізу даних про державні закупівлі та фінансові операції, проте системне застосування ШІ в антикорупційній діяльності ще не налагоджено.

Застосування ШІ у протидії корупції має ґрунтуватися на чітких правових та етичних засадах. Правове регулювання повинно забезпечувати законність,

прозорість та підзвітність використання ШІ, а також захист прав людини. Основною перешкодою у системному впровадженні цифрових інструментів протидії корупції в економічних відносинах на основі ШІ в Україні є відсутність необхідної законодавчої основи для системного застосування цієї технології. Як наслідок, залишаються неврегульованим питання правових підстав, умов і процедур застосування ШІ у протидії корупції, вимоги до обробки інформації з використанням ШІ, вимоги до моделей цієї технології, які можуть інтегруватися до національного антикорупційного механізму.

На нашу думку, умови, на основі яких можуть розгортатися цифрові інструменти на основі ШІ, призначені для запобігання корупції повинні включати такі вимоги, як точність і пояснюваність моделей ШІ, дотримання етичних норм, відповідність нормам права, відкритість і аудит. У цьому дискурсі новим етапом є поява моделей так званих розмірковуючих моделей (reasoning) [38]. Одним із нашумілих прикладів є сімейство великих мовних моделей (LLM) Deepseek-R1 [39]. Сюди входить донавчена генеративна модель Deepseek-V3 із 671 мільярдом параметрів, а також серія значно менших дистильованих [40] мовних моделей від різних виробників із тим же функціоналом обмірковування [41]. Для короткої довідки зазначимо, що дані мовні моделі мають більшу точність порівняно зі звичайними, неміркуючими, за рахунок огляду проблематики [42] одержаного ними запиту (prompt) [43]. Себто перед наданням відповіді шляхом звичайного передбачення наступних найбільш ймовірних токенів ці мовні моделі спершу осмислюють суть питання та можливі підходи до вирішення проблеми (що можна бачити в їхніх термінальних або графічних інтерфейсах, зазвичай між тегами <think></think>) [44]. Таким чином, висновки генеративних моделей відтепер можливо перевірити не лише за підсумками остаточної відповіді, а й побачити їхнє власне обґрунтування, що особливо важливо в рамках людиноорієнтованих рішень щодо протидії корупції.

На відміну від України, в Євросоюзі 13 березня 2024 року було прийнято Регламент (ЄС) 2024/1689 Європарламенту та Ради, який встановлює гармонізовані правила щодо штучного інтелекту та вносить зміни до регламентів

(ЄС) № 300/2008, (ЄС) № 167/2013, (ЄС) № 168 /2013, (ЄС) 2018/858, (ЄС) 2018/1139, (ЄС) 2019/2144 і директив 2014/90/ЄС, (ЄС) 2016/797 та (ЄС) 2020/1828 (Акт про штучний інтелект) [45]. Згаданий законодавчий акт Європейського парламенту визначає загальні принципи, умови застосування ШІ у різних сферах суспільного життя, передбачає гарантії дотримання прав споживачів контенту, згенерованого повністю або з використанням ШІ, запроваджує механізм відповідальності за правопорушення, вчинені з використанням ШІ, тощо.

Нормами Акту про ШІ передбачено диференційоване законодавче регулювання використання систем ШІ залежно від так званої «зони ризику»: неприйнятний ризик, високий ризик, обмежений ризик та мінімальний ризик. Залежно від «зони ризику», у розділі ІІ Закону викладено заборонені практики використання систем ШІ на території ЄС.

Підготовка до ухвалення вказаного Регламенту ЄС та вироблення уніфікованих підходів до унормування використання ШІ вирізнялися комплексністю й ґрунтувалися на низці актів рекомендаційного характеру, прийнятих керівними інституціями ЄС, серед яких: Резолюція Європейського Парламенту з рекомендаціями Комісії з норм цивільного права щодо робототехніки 2015/2103(INL) [46], Повідомлення Європейської Комісії до Європейського Парламенту, Європейської Ради, Ради Європейського економічного та соціального Комітету та Комітету регіонів «Штучний інтелект для Європи» [47], Рекомендації Експертної групи високого рівня зі штучного інтелекту з етики для надійного ШІ [48], Рекомендація Комітету міністрів Ради Європи Cm/rec (2020) державам-членам щодо впливу алгоритмічних систем на права людини [49], «Біла книга зі штучного інтелекту: європейський підхід до досконалості і довіри» [50], Звіт Європейської Комісії про наслідки для безпеки і відповідальності ШІ, Інтернету речей та робототехніки [51].

Враховуючи потенційні ризики та можливі негативні впливи використання ШІ у площині гарантування й забезпечення основних прав людини та етичних норм, всі ці документи містять пропозиції щодо їх дотримання. Зокрема,

Резолюція Європейського Парламенту 2015/2103 (INL) стала підґрунтям для запровадження єдиних правових стандартів і розробки конкретних правових інструментів використання ШІ на території ЄС. У ній передбачені засади неухильного дотримання законодавства про захист персональних даних та способи подолання проблем залежності людини від технічних приладів. У Рекомендаціях Експертної групи високого рівня з ШІ з етики для надійного ШІ також визначені основні принципи його використання: орієнтованість на людину, тобто системи ШІ повинні розроблятися, використовуватися й контролюватися, забезпечуючи дотримання основних прав людини та етичних норм. «Надійний ШІ» вимагає, щоб алгоритми були законними, безпечними, надійними та досить стійкими, щоб справлятися з помилками або невідповідностями на всіх етапах циклу системи ШІ. Рекомендації містять 7 ключових вимог, яким мають відповідати системи ШІ, щоб вважатися надійними: 1) людські права та нагляд; 2) технічна надійність і безпека; 3) конфіденційність й управління даними; 4) прозорість; 5) різноманітність, відсутність дискримінації та справедливість; 6) суспільне та екологічне благополуччя; 7) підзвітність. Рекомендація Комітету Міністрів Ради Європи Cm/rec (2020) державам-членам щодо впливу алгоритмічних систем на права людини містить зобов'язання держав стосовно захисту й заохочення прав людини та основних свобод у контексті алгоритмічних систем.

Базуючись на цих напрацюваннях, положеннями Закону встановлено єдині зобов'язання для операторів і гарантії однакового захисту суспільного інтересу та прав осіб на всьому внутрішньому ринку на основі статті 114 Договору про функціонування ЄС. Метою Закону є покращення функціонування внутрішнього ринку ЄС та сприяння використанню орієнтованого на людину й надійного ШІ, за умови одночасного забезпечення високого рівня захисту здоров'я, безпеки та основних прав, закріплених у Хартії основних прав Європейського Союзу, включаючи демократію, верховенство права, захист навколишнього середовища та підтримку інновацій, протидію шкідливому впливу систем ШІ (стаття 1).

Отже, підхід, який застосовується в Акті «Про штучний інтелект» стосовно прав людини при використанні систем ШІ, виявляється в тому, що ці системи не можуть створювати більш істотні ризики для людини порівняно з традиційними підходами гарантування її прав та свобод, які реалізуються у традиційний спосіб.

Прийняття Європейським Парламентом уніфікованого, комплексного законодавчого акту, обов'язкового для усіх держав-членів ЄС фактично стало підґрунтям для розробки та прийняття перших міжнародних договорів зобов'язального характеру у цій сфері правового регулювання.

Першим міжнародним юридично обов'язковим договором, що регулює ШІ стала Рамкова Конвенція Ради Європи "Про права людини, штучний інтелект, демократію та верховенство права" [52]. Вона була відкрита для підписання 5 вересня 2024 року у Вільнюсі та підписана низкою держав, серед яких Велика Британія, США, країни-члени ЄС та інші.

Метою цієї Конвенції є забезпечення дотримання прав людини, принципів демократії та верховенства права у всьому життєвому циклі систем ШІ.

Документ встановлює загальні принципи та зобов'язання, які можуть бути деталізовані подальшими нормативними актами.

Конвенція містить низку ключових зобов'язань для держав-учасниць:

- Обов'язки держав щодо регулювання ШІ: встановлення належного законодавчого та адміністративного регулювання відповідно до рівня ризиків, пов'язаних із ШІ.
- Принципи відповідального використання ШІ: прозорість, підзвітність, недискримінація, захист приватності та безпека алгоритмічних систем.
- Захист демократичних інституцій: запобігання використанню ШІ для маніпуляцій громадською думкою, цензури або незаконного стеження.
- Право на оскарження рішень ШІ: користувачі повинні мати доступ до інформації про використання ШІ та механізмів оскарження рішень, прийнятих системами ШІ.

- Нагляд та міжнародна співпраця: створення незалежного наглядового механізму в кожній країні та обмін досвідом між державами-учасницями.

Поряд із зазначеними положеннями, Конвенція також встановлює важливі правила набрання нею чинності. Йдеться, зокрема, про необхідність її ратифікації щонайменше п’ятьма державами, три з яких повинні бути державами-членами Ради Європи та здійснення обов’язкової гармонізації норм національного законодавства з нормами цього міжнародно-правового акту.

Для України, як держави-члена Ради Європи проаналізована Конвенція є відкритою для підписання.

На нашу думку, ратифікація Конвенції і подальше виконання Україною зобов’язань щодо формування ефективного законодавчого підґрунтя для розробки і застосування ШІ в різних сферах суспільного життя буде сприяти посиленню правових гарантій прав учасників інформаційних відносин у цій сфері правового регулювання. З інших правових позицій, Конвенція фактично встановлює узагальнені правові підходи до розробки національного законодавства про ШІ, визначає основні види негативних зобов’язань держав-учасниць Конвенції щодо необхідності запровадження системи гарантій недопущення вчинення правопорушень. Це є особливо важливим в умовах розробки національної великої мовної моделі (LLM), що анонсував Міністр цифрової трансформації Михайло Федоров [53], використання цифрового ШІ-аватара Вікторії у якості представниці МЗС з консульських питань [54] та чималої кількості розробок на поприщі військової справи (дрони, розмінування тощо) [55]. По суті, технології слабкого штучного інтелекту будуть оточувати громадян в усіх сферах їхнього життя, відтак правове регулювання, зокрема і в контексті протидії корупції, стане необхідним кроком для нормалізації взаємодії людина-технологія.

Враховуючи євроінтеграційний курс України, імплементація в норми національного законодавства положень Акту про ШІ, може забезпечити комплексне врегулювання механізму інтеграції ШІ у механізм протидії корупції

з урахуванням цілей сталого економічного розвитку нашої держави. Водночас, поряд із законодавчим забезпеченням врегулювання питань, порушених у цій роботі, застосування ШІ у протидії корупції обумовлює необхідність вирішення ще окремих системних проблем, які потребують свого вирішення.

Так, застосування ШІ у протидії корупції має забезпечувати баланс між правом на інформацію та правом на конфіденційність даних. З одного боку, доступ до інформації є необхідною умовою для ефективного виявлення та розслідування корупційних правопорушень. З іншого боку, необхідно забезпечити захист персональних даних та приватності осіб, щодо яких проводиться перевірка.

Для забезпечення балансу між цими правами необхідно застосовувати такі заходи, як анонімізація та псевдонімізація даних, автоматизовані рішення з переглядом людиною, а також чіткі процедури обробки та захисту даних відповідно до норм національного законодавства України про захист персональних даних. Відтак визначальними будуть і підходи відкритого програмного забезпечення (open source). Міністерство цифрової трансформації вже рухається в цьому напрямку, прикладом чого може слугувати відкриття сирцевого коду їхнього чи не головного програмного продукту - додатку “Дія”[56]. Оскільки відкритість програмної логіки (backend) аж ніяк не означає легкодоступність даних, а навпаки - передбачає можливість усіх охочих долучитися до пошуків прогалин у безпеці (алгоритмах шифрування, способах одержання та обробки даних тощо) та переконатися в надійності системи, якщо недоліків не знайдено. Відтак подібні розробки перебувають під постійним наглядом спільноти відкритого ПЗ (Collaborative participation and shared responsibility) [57].

Кроки в цьому напрямку здійснюються не лише в Україні, а і світі, адже помітним є вихід чималої кількості відкритих моделей ШІ, а саме:

- тривала розробка та поширеність Stable Diffusion від Stability AI [58]
- аналогічне можна сказати про Llama від Meta [59]
- вищезгадане сімейство моделей Deepseek-R1 [60]

- інша азіатська розробка, а саме моделі Qwen від Alibaba Group [61]

- нещодавній анонс першої з часів GPT-2 відкритої моделі від OpenAI [62]

Насамкінець зазначимо ряд можливих морально-етичних дилем пов'язаних із використанням ШІ, а саме:

- відсутність емпатії (адже ШІ не обов'язково володіти почуттями та розумінням про емоції для аналізу даних на предмет корупційної діяльності, поясненням для цього можуть слугувати уявні експерименти "Кімната Марії" [63] та "Китайська кімната" [64])

- відсутність або розмитість меж відповідальності (оскільки наразі не є юридично окресленим та відсутнє прецедентне право щодо сфер відповідальності ШІ, людини-оператора та розробника, прикладом ситуації може слугувати перший випадок зі збиття людини AI-керованим авто у 2018 році [65], а іншою важливою складовою для гіпотетичних ситуацій є "проблема вагонетки" в якій AI доведеться обирати смерть тих чи інших людей, коли інший вибір не буде можливий [66])

- проблема навчання, а згодом і визначення новизни авторства (адже збір тренувальних даних не завжди супроводжується офіційним дозволом на використання цієї інформації, і ця проблема піднята у серії сучасних протестів митців [67] та кінематографістів [68] проти AI-продукції)

Усі ці питання дедалі частіше поставатимуть перед законодавцями як України, так і світу, у галузі боротьби з корупційною діяльністю зокрема. Відтак хочемо наголосити, що необхідно чітко визначитися з понятійним апаратом, опрацювати всі гіпотетичні ситуації до їхнього втілення в життя (а за можливості, уникнення тих ситуацій, котрі можуть принижувати людську честь та гідність, а особливо стати причиною осуду, каліцтва тощо). І, як остаточна рекомендація, це приділити особливу увагу відкритості ШІ-алгоритмів, адже без розуміння внутрішньої будови та процесів всередині штучного інтелекту людство як таке може потрапити в пастку втрати самостійності, неправомірних рішень та/або гірших сценаріїв.

**Висновки.** Проведене дослідження дозволяє констатувати, що застосування ШІ має значний потенціал для підвищення ефективності протидії корупції в економіці. Проте, для реалізації цього потенціалу необхідно забезпечити належне правове та етичне регулювання використання ШІ, а також захист прав людини.

Пропозиції щодо вдосконалення законодавства включають:

- Імплементацію Регламенту ЄС про ШІ у національне законодавство.
- Внесення окремих змін до Закону України «Про запобігання корупції» в частині формування правових засад інтеграції ШІ у механізм запобігання корупційним правопорушенням.
- Затвердження Етичного кодексу застосування ШІ для протидії корупції.
- Забезпечення балансу між правом на інформацію та правом на конфіденційність даних.

**Список джерел**

1. Transparency International. (2023). *Corruption Perceptions Index 2023*. https://www.transparency.org/en/cpi/2023.

2. UNODC. (2003). *United Nations Convention against Corruption*. https://www.unodc.org/unodc/en/corruption/uncac.html.

3. Генеральна Асамблея ООН. (1948). *Загальна декларація прав людини* (Резолюція 217 А (III)). Відомості Верховної Ради України, 1993, № 42, ст. 409.

4. Рада Європи. (1950). *Конвенція про захист прав людини і основоположних свобод*. https://zakon.rada.gov.ua/laws/show/995_004.

5. Верховна Рада України. (2014). *Про запобігання корупції* (Закон № 1700-VII). https://zakon.rada.gov.ua/laws/show/1700-18.

6. Zinnbauer, D. (2025). *Artificial intelligence in anti-corruption: A timely update on AI technology* (U4 Brief 2025:1). Chr. Michelsen Institute. https://www.cmi.no/publications/file/9480-artificial-intelligence-in-anti-corruption-a-timely-update-on-ai-technology.pdf.

7. So, S. J. (2025, January 10). *Artificial intelligence in anti-corruption: Opportunities and challenges*. Corruption Watch South Africa. https://www.corruptionwatch.org.za/artificial-intelligence-in-anti-corruption-opportunities-and-challenges/.

8. Заярний, О., & Ілларіонова, С. (2025). Застосування технологій штучного інтелекту для навчання майбутніх працівників на прикладі CHATGPT: педагогічний, правовий і технологічний аспекти. *Інформаційні технології та засоби навчання, 105*(1), 158–177. https://doi.org/10.33407/itlt.v105i1.5835.

9. ISO. (2020). *Information technology - Artificial intelligence - Overview of trustworthiness in artificial intelligence* (ISO/IEC TR 24028:2020). https://www.iso.org/obp/ui/#iso:std:isoiec:tr:24028:ed-1:v1:en.

10. IBM. (2024). *What is artificial general intelligence (AGI)?* https://www.ibm.com/think/topics/artificial-general-intelligence.

11. European Commission. (2023). *Building AI-powered government: The Estonian experience*. https://reform-support.ec.europa.eu/document/download/27b74e18-8a00-4cb5-84ca-9ede00a63e63_en.

12. e-Estonia. (2024). *AI use in public sector is a team sport*. https://e-estonia.com/ai-use-in-public-sector-is-a-team-sport/.

13. Міністерство цифрової трансформації України. (n.d.). *Державні послуги онлайн. Дія*. https://diia.gov.ua/faq/14.

14. Euronews. (2025, February 25). *Estonia announces new AI initiative to integrate chatbots into schools*. https://www.euronews.com/next/2025/02/25/estonia-announces-new-ai-initiative-with-openai-to-roll-out-chatbots-in-schools.

15. Taxadmin AI. (n.d.). *Denmark - AI Country Report - AI Tax Administration*. https://taxadmin.ai/country/denmark-ai-country-report/.

16. Darktrace. (2024). *Danish Tax Authority uses Darktrace for cybersecurity*. https://www.darktrace.com/news/danish-tax-authority-secures-critical-infrastructure-with-darktrace-ai.

17. Avalara, Inc. (2020, June). *Can artificial intelligence curb tax evasion?* https://www.avalara.com/blog/en/apac/2020/06/can-artificial-intelligence-curb-tax-evasion.html.

18. AML Intelligence. (2024, October). *NEWS: Germany's FIU examining how it can use AI to fight financial crime*. https://www.amlintelligence.com/2024/10/news-germanys-fiu-examining-how-it-can-use-ai-to-fight-financial-crime/.

19. BMBF. (2024). *Artificial Intelligence*. Bundesministerium für Bildung und Forschung. https://www.bmbf.de/EN/Research/EmergingTechnologies/ArtificialIntelligence/artificialintelligence_node.html.

20. Організація економічного співробітництва та розвитку (OECP). (2020). *Використання цифрових технологій для боротьби з корупцією*. OECD Publishing. https://www.oecd-ilibrary.org/governance/using-digital-technologies-to-combat-corruption_99738734-en.

21. AI Watch. (2021). *France: Public Sector dimension of AI Strategy*. European Commission. https://ai-watch.ec.europa.eu/topics/public-sector/public-sector-dimension-ai-national-strategies/france-public-sector-dimension-ai-strategy_en.

22. UNDP. (2024). *Pairing with AI for Public Sector Impact in Singapore*. https://www.undp.org/policy-centre/singapore/blog/pairing-ai-public-sector-impact-singapore.

23. National Research Foundation, Singapore. (n.d.). *Smart Nation and Digital Economy*. https://www.nrf.gov.sg/rie-ecosystem/ecosystem-wide-fis/snde-fis/

24. AI Singapore. (n.d.). *Home*. https://aisingapore.org/.

25. SAS. (n.d.). *Brazil tackles insurance fraud with AI and analytics*. https://www.sas.com/tr_tr/customers/cnseg.html.

26. Digiampietr, L. A., Norton, T., & Meira, L. (2008). Uses of artificial intelligence in the Brazilian customs fraud detection system. *The Proceedings of the 9th Annual International Digital Government Research Conference*, 158–177. https://doi.org/10.1145/1367832.1367864.

27. OECD AI Policy Observatory. (2025, March 25). *AI-driven fraud sparks financial losses in Brazil*. https://oecd.ai/en/incidents/2025-03-25-3fd0.

28. European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council (Artificial Intelligence Act)*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

29. OpenAI. (2022). *Introducing ChatGPT*. https://openai.com/index/chatgpt/.

30. OpenAI. (2022). *Introducing Whisper*. https://openai.com/index/whisper/.

31. Google. (2023). *Introducing Gemini: Our largest and most capable AI model*. https://blog.google/technology/ai/google-gemini-ai/.

32. Google. (2023). *Introducing Gemma 3: The most capable model you can run on a single GPU or TPU*. https://blog.google/technology/developers/gemma-3/.

33. Google Research. (2017). *Transformer: A novel neural network architecture for language understanding*. https://research.google/blog/transformer-a-novel-neural-network-architecture-for-language-understanding/.

34. Microsoft. (n.d.). *Microsoft Copilot: Ihr KI-Begleiter*. https://copilot.microsoft.com/.

35. Microsoft Community Hub. (2024). *Introducing Phi-4: Microsoft's newest small language model specializing in complex reasoning*. https://techcommunity.microsoft.com/blog/aiplatformblog/introducing-phi-4-microsoft%E2%80%99s-newest-small-language-model-specializing-in-comple/4357090.

36. Meta AI. (2025). *The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation*. https://ai.meta.com/blog/llama-4-multimodal-intelligence/.

37. OpenAI. (2022). *Announcing The Stargate Project*. https://openai.com/index/announcing-the-stargate-project/.

38. Raschka, S. (2025). *Understanding reasoning LLMs*. https://sebastianraschka.com/blog/2025/understanding-reasoning-llms.html.

39. DeepSeek. (2025). *DeepSeek-R1 Release*. https://api-docs.deepseek.com/news/news250120.

40. Google for Developers. (2025). *Великі мовні моделі: донавчання, дистиляція і інженерія запитів*. https://developers.google.com/machine-learning/crash-course/llm/tuning.

41. Ollama. (n.d.). *Deepseek-r1*. https://ollama.com/library/deepseek-r1.

42. DigitalOcean. (2023). *Understanding reasoning in Large Language Models: Overview of the paper "Towards Reasoning in Large Language Models: A Survey"*. https://www.digitalocean.com/community/tutorials/understanding-reasoning-in-llms.

43. Google Cloud. (n.d.). *Prompt Engineering for AI Guide*. https://cloud.google.com/discover/what-is-prompt-engineering.

44. Towards Data Science. (2025). *How to train LLMs to "think" (o1 & DeepSeek-R1)*. https://towardsdatascience.com/how-to-train-llms-to-think-o1-deepseek-r1/.

45. European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council (Artificial Intelligence Act)*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

46. European Parliament. (2017). *Civil Law Rules on Robotics (2015/2103(INL))*. https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.htm.

47. European Commission. (2018). *Communication: Artificial Intelligence for Europe*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN.

48. AI HLEG. (2019). *Ethics guidelines for a trustworthy AI*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

49. Committee of Ministers. (2020). *Recommendation CM/Rec(2020)1 on the human rights impacts of algorithmic systems*. Council of Europe. https://www.coe.int/en/web/freedom-expression/committee-of-ministers-adopted-texts/-/asset_publisher/aDXmrol0vvsU/content/recommendation-cm-rec-2020-1-of-the-committee-of-ministers-to-member-states-on-the-human-rights-impacts-of-algorithmic-systems.

50. European Commission. (2020). *White Paper on Artificial Intelligence – A European approach to excellence and trust*. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.

51. European Commission. (2020). *Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics*. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52020DC0064.

52. Рада Європи. (2024). *Рамкова Конвенція Ради Європи «Про штучний інтелект, права людини, демократію та верховенство права»*. https://rm.coe.int/-1493-10-1b-committee-on-artificial-intelligence-cai-b-draft-framework/1680aee411.

53. Міністерство цифрової трансформації України. (2025, Січень). *Михайло Федоров про розвиток ШІ в Україні та роботу над національною LLM*. https://thedigital.gov.ua/news/mikhaylo-fedorov-pro-rozvitok-shi-v-ukraini-ta-robotu-nad-natsionalnoyu-llm.

54. МЗС України. (2024). *Вітальне слово представниці МЗС з консульських питань Вікторії Ші*. https://mfa.gov.ua/news/vitalne-slovo-predstavnici-mzs-z-konsulskih-pitan-viktoriyi-shi.

55. АрміяInform. (2025, Лютий 11). *Штучний інтелект допоможе Міноборони України з розмінуванням*. https://armyinform.com.ua/2025/02/11/shtuchnyj-intelekt-dopomozhe-minoborony-ukrayiny-z-rozminuvannyam/.

56. Дія. (n.d.). *Дія.Open Source*. https://opensource.diia.gov.ua/.

57. Red Hat. (2024). *Open source culture: 9 core principles and values*. https://www.redhat.com/en/blog/open-source-culture-9-core-principles-and-values

58. Stability AI. (n.d.). *A model for generating professional-grade images*. https://stability.ai/.

59. Meta. (n.d.). *Llama 4: Leading intelligence. Unrivaled speed and efficiency*. https://www.llama.com/.

60. DeepSeek-AI. (2025). *DeepSeek-R1* [Source code]. GitHub. https://github.com/deepseek-ai/DeepSeek-R1

61. Alibaba Cloud. (2025). *Qwen LLMs*. https://www.alibabacloud.com/help/en/model-studio/what-is-qwen-llm.

62. Reuters. (2025, March 31). *OpenAI plans to release open-weight language model in coming months*. https://www.reuters.com/technology/artificial-intelligence/openai-plans-release-open-weight-language-model-coming-months-2025-03-31/.

63. Jackson, F. (1986). What Mary didn't know. *The Journal of Philosophy, 83*(5), 291–295. https://web.ics.purdue.edu/~drkelly/JacksonWhatMaryDidntKnow1986.pdf.

64. Searle, J. R. (1990). Is the brain's mind a computer program? *Scientific American, 262*(1), 26–31. https://www.cs.princeton.edu/courses/archive/spr06/cos116/Is_The_Brains_Mind_A_Computer_Program.pdf.

65. The New York Times. (2018, March 20). *How a self-driving Uber killed a pedestrian in Arizona*. https://www.nytimes.com/interactive/2018/03/20/us/self-driving-uber-pedestrian-killed.html.

66. Britannica, T. Editors of Encyclopaedia. (2021). *Trolley problem*. Encyclopedia Britannica. https://www.britannica.com/topic/trolley-problem.

67. VICE. (2022, December). *Artists are revolting against AI art on ArtStation*. https://www.vice.com/en/article/artists-are-revolt-against-ai-art-on-artstation/.

68. The Guardian. (2023, October 1). *How Hollywood writers triumphed over AI – and why it matters*. https://www.theguardian.com/culture/2023/oct/01/hollywood-writers-strike-artificial-intelligence.

* * *