УДК 347.772.3
DOI https://doi.org/10.32782/TNU-2707-0581/2026.5/14

Фірман Д. І.
https://orcid.org/0000-0001-7598-6394
Львівський національний університет імені Івана Франка

БАЛАНС МІЖ ДОКТРИНОЮ ДОБРОСОВІСНОГО ВИКОРИСТАННЯ (FAIR USE) ТА ЛІЦЕНЗУВАННЯМ У КОНТЕКСТІ НАВЧАННЯ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ: ПРОБЛЕМАТИКА ПРАВОМІРНОСТІ ВИКОРИСТАННЯ ОБ'ЄКТІВ АВТОРСЬКОГО ПРАВА

Стаття присвячена правовому аналізу використання об'єктів авторського права у процесі навчання моделей штучного інтелекту та визначенню критеріїв правомірності такого використання. Встановлено, що поширення генеративного ШІ та необхідність використання значних масивів даних для навчання відповідних моделей загострюють питання забезпечення прав та законних інтересів авторів і правовласників.

Проаналізовано сучасні підходи до регулювання використання охоронюваних творів для навчання ШІ у Європейському Союзі та США, зокрема положення Регламенту (ЄС) 2024/1689 (AI Act), винятки щодо глибинного аналізу текстів і даних та доктрину fair use. Звернуто увагу на транскордонний характер використання навчальних даних, що ускладнює визначення застосовного права та юрисдикції у спорах щодо порушення авторських прав.

Встановлено, що правомірність використання творів для навчання моделей ШІ не може визначатися виключно фактом їх доступності в мережі Інтернет. При її оцінці мають враховуватися правова підстава доступу до твору, характер та обсяг його використання, наявність дозволу або застереження правовласника, можливість застосування винятків і обмежень авторського права, здатність моделі відтворювати охоронювані елементи твору, а також вплив такого використання на економічні інтереси правовласника. Звернуто увагу на доцільність застосування правовласниками технічних засобів і спеціальних позначень для обмеження автоматизованого збору та використання їхнього контенту для навчання моделей ШІ.

Окрему увагу приділено механізму opt-out та проблемам його практичного застосування, зокрема недостатній обізнаності правовласників щодо можливості застереження від використання їхніх творів для навчання ШІ. Досліджено сучасну судову практику та встановлено складність доведення факту використання конкретних творів у навчальних масивах і характеру такого використання.

Ключові слова: *авторське право, штучний інтелект, навчання моделей штучного інтелекту, fair use (добросовісне використання), ліцензування, opt-out, цифрове середовище.*

Постановка проблеми. Розвиток генеративного штучного інтелекту поставив перед авторським правом питання щодо визначення допустимих меж використання охоронюваних творів для навчання моделей ШІ. Масштаб такого використання ускладнює застосування традиційної моделі індивідуального ліцензування, оскільки навчальні масиви можуть охоплювати значну кількість творів, встановлення правового статусу та правовласників яких потребує істотних організаційних і фінансових ресурсів.

Водночас сучасне правове регулювання поступово формує механізми узгодження інтересів правовласників і розробників ШІ. У праві ЄС важливе значення мають винятки щодо глибинного аналізу текстів і даних, передбачені Директивою (ЄС) 2019/790, у поєднанні з положеннями Регламенту (ЄС) 2024/1689 (AI Act).

Разом із тим наявність спеціального регулювання не усуває питання щодо правової підстави використання конкретних об'єктів авторського права під час навчання ШІ. Відкритий доступ до твору

© Фірман Д. І., 2026
Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0

сам по собі не означає можливості його вільного використання, а навчальні масиви можуть містити контент, первинне розміщення якого здійснено без дозволу правовласника. Відповідно, актуальними залишаються питання дотримання виключних прав автора, застосування передбачених законом винятків та ефективності ліцензійних механізмів.

Особливого значення ця проблематика набуває при зіставленні різних правових підходів. Якщо право США допускає оцінку використання творів для навчання ШІ крізь призму доктрини fair use, то право ЄС виходить із системи спеціально визначених винятків та обмежень авторського права. За таких умов основна проблема полягає не у відсутності правового регулювання, а у визначенні оптимального співвідношення між дозволеним законом використанням творів та їх ліцензуванням, яке забезпечувало б належну охорону інтересів правовласників, правову визначеність для розробників ШІ та не створювало необґрунтованих перешкод для технологічного розвитку.

Аналіз останніх досліджень і публікацій. Проблематика правомірності використання об'єктів авторського права у процесі навчання моделей штучного інтелекту протягом останніх років перебуває у центрі уваги як науковців і практикуючих юристів, так і міжнародних та європейських інституцій. Окремі аспекти цієї проблематики досліджуються Європейським офісом з питань штучного інтелекту (AI Office), Європейською радою зі штучного інтелекту (AI Board), Комітетом з питань комунікацій та цифрових технологій Палати лордів Великої Британії. Значну увагу питанням взаємодії штучного інтелекту та права інтелектуальної власності приділяє Всесвітня організація інтелектуальної власності (ВОІВ), у межах діяльності якої проводяться міжнародні дискусії та конференції, присвячені актуальним технологічним викликам для існуючої системи правової охорони інтелектуальної власності. Додаткового значення ця проблематика набула у процесі розроблення та прийняття Регламенту (ЄС) 2024/1689 про штучний інтелект, положення якого, серед іншого, безпосередньо стосуються дотримання законодавства ЄС у сфері авторського права провайдерами моделей ШІ загального призначення.

Питання взаємодії авторського права та технологій штучного інтелекту, зокрема використання охоронюваних творів для навчання моделей ШІ, висвітлювали у своїх працях К. Зеров, Ц. Міллер, М. Ділворт, Тітінтін Веня, Ф. Ріксон, О. Кульчій, А. С. Бугаєць, С. Глотов та інші.

Водночас особливого значення для дослідження набуває правозастосовна практика. Судові спори щодо використання охоронюваних творів для навчання моделей ШІ, а також позиції органів у сфері інтелектуальної власності дають змогу простежити формування практичних підходів до визначення меж правомірного використання творів.

Постановка завдання. Метою дослідження є аналіз правових засад використання об'єктів авторського права у процесі навчання моделей штучного інтелекту, визначення критеріїв правомірності такого використання, дослідження співвідношення доктрини fair use та ліцензійних механізмів у контексті забезпечення балансу між правами та законними інтересами правовласників і потребами технологічного розвитку.

Виклад основного матеріалу. Штучний інтелект за відносно короткий період трансформувався з інноваційної технології у повсякденний інструмент, який активно використовується як бізнесом, так і пересічними користувачами. За допомогою систем ШІ щодня створюються значні обсяги текстового, графічного, аудіовізуального та іншого контенту. Зокрема, за даними Міністерства цифрової трансформації України, у 2025 році технологіями штучного інтелекту користувалися 42% дорослого населення та 70% підлітків [1]. Зростання попиту на такі технології стимулює розробників удосконалювати існуючі та створювати нові моделі ШІ, функціонування яких потребує значних обсягів даних.

Водночас здатність генеративних систем ШІ створювати новий контент безпосередньо залежить від процесу їх попереднього навчання. Зокрема, OpenAI зазначає, що моделі, які забезпечують роботу ChatGPT, тренуються з використанням загальнодоступної інформації в Інтернеті, даних, отриманих від третіх осіб, а також інформації, наданої або створеної користувачами, тренерами та дослідниками [2]. Однак, безсумнівно, сам факт відкритого доступу до інформації не означає відсутності правової охорони такого контенту. Більшість творів, розміщених у мережі Інтернет, продовжують охоронятися авторським правом незалежно від того, чи є доступ до них вільним. Відтак виникає питання щодо допустимості використання таких творів для навчання моделей штучного інтелекту без отримання згоди авторів або інших правовласників. Значна частина суб'єктів авторського права може не погоджуватися з використанням належних їм творів для навчання моделей штучного інтелекту, особливо якщо таке використання здійснюється без їх

попереднього повідомлення, згоди чи виплати винагороди. Водночас відсутність явно вираженої незгоди правовласника не завжди означає, що він фактично погоджується на використання свого твору для таких цілей.

Особливого значення це питання набуває в контексті передбаченого правом ЄС механізму застереження прав (opt-out). Відповідно до ст. 4 Директиви (ЄС) 2019/790 допускається відтворення та витягування правомірно доступних творів та інших об'єктів для цілей глибинного аналізу текстів та даних, якщо відповідні права не були прямо застережені їх правовласниками [3]. Щодо контенту, розміщеного у відкритому доступі в мережі Інтернет, Директива передбачає можливість здійснення такого застереження у машинозчитуваний спосіб.

На практиці реалізація цього механізму породжує низку питань. Автор або інший правовласник може не знати про необхідність здійснення відповідного застереження, не володіти технічними засобами для його належного оформлення або взагалі не мати можливості самостійно встановити відповідні технічні позначення, якщо твір поширюється через сторонню платформу. Не завжди очевидним для правовласника є і те, що розміщення твору у відкритому доступі потенційно створює можливість його використання для навчання моделей ШІ за відсутності належного застереження прав.

Вважаю, що сама лише відсутність технічно оформленого застереження не повинна автоматично ототожнюватися зі згодою правовласника на використання твору для навчання моделі ШІ. Особливо проблемним такий підхід є у випадках, коли правовласник фактично не має можливості контролювати технічний спосіб поширення твору на сторонній платформі. Відтак ефективність моделі opt-out залежить не лише від її нормативного закріплення, а й від доступності та зрозумілості способів застереження прав, їх технічної стандартизації та належного врахування таких застережень розробниками моделей ШІ.

Крім того, хоча ст. 53(1)(d) Регламенту (ЄС) 2024/1689 про штучний інтелект встановлює обов'язок провайдерів моделей ШІ загального призначення складати та оприлюднювати достатньо детальне резюме контенту, використаного для навчання моделі [4], практична реалізація цієї вимоги залишається недостатньо ефективною. Результати дослідження, проведеного вченими Триніті-коледжу Дубліна, свідчать, що на практиці відповідні резюме не завжди оприлюднюються або їх складно знайти [5]. Однією з причин є відсутність централізованого механізму їх оприлюднення: Регламент не визначає єдиного місця, де провайдери повинні розміщувати такі резюме, фактично залишаючи вибір способу та місця публікації за ними. За таких умов сам факт нормативного закріплення обов'язку щодо прозорості ще не забезпечує правовласникам простого та ефективного доступу до інформації, необхідної для встановлення факту можливого використання їхніх творів під час навчання моделей ШІ.

Додаткове занепокоєння викликає і те, що дані, зібрані з великої кількості відкритих джерел, потенційно можуть містити контент, який був розміщений в Інтернеті з порушенням авторських прав або без належних правових підстав для його подальшого використання. За таких умов виникають складнощі з перевіркою правового статусу кожного окремого об'єкта, що включається до навчальних наборів даних, а також із визначенням відповідальності за можливе використання неправомірно оприлюднених творів. Крім того, OpenAI, як і інші розробники не розкривають детальної інформації щодо складу та походження матеріалів, які надаються тренерами моделей або використовуються для їх подальшого вдосконалення, вважаючи цю інформацію важливою частиною своєї конкурентної переваги.

У цьому контексті заслуговують на увагу висновки Звіту Комітету з питань комунікацій та цифрових технологій Великої Британії за 2024–2026 роки, в якому зазначається, що широке неліцензійне використання охоронюваних творів у поєднанні з обмеженою прозорістю розробників штучного інтелекту щодо навчальних даних їхніх моделей суттєво ускладнює реалізацію прав авторів та інших правовласників [6]. Наслідком цього є фактична неможливість для багатьох суб'єктів авторського права перевірити, чи використовувалися їхні твори під час навчання моделей, а також належним чином захистити свої права у разі виявлення такого використання.

У США правомірність використання охоронюваних творів без дозволу правовласника за певних умов може обґрунтовуватися доктриною fair use (добросовісного використання), закріпленою у параграфі 107 Закону США про авторське право [7]. Її метою є забезпечення балансу між виключними правами автора та суспільним інтересом у розвитку науки, освіти й технологій. Саме на fair use розробники моделей ШІ часто посилаються як на правову підставу використання творів для навчання.

Загалом використання об'єктів авторського права для навчання моделей ШІ може безпосередньо впливати на економічні інтереси авторів/правовласників, зокрема у випадках, коли створений за допомогою ШІ контент конкурує з результатами людської творчої діяльності або зменшує попит на них. Водночас встановлення надмірно обтяжливих вимог щодо використання кожного окремого твору для навчання моделей ШІ може істотно ускладнити розвиток відповідних технологій та негативно позначитися на суміжних сферах економіки. Тому правове регулювання у цій сфері має бути спрямоване на пошук балансу між забезпеченням майнових інтересів авторів і правовласників та збереженням можливостей для технологічного розвитку.

У цьому контексті заслуговує на увагу пропозиція О. О. Кульчія щодо запровадження механізму, подібного до обов'язкового або розширеного колективного управління авторськими правами [8, с.143]. Його сутність полягає у створенні правового механізму, який дозволяв би використовувати публічно доступні об'єкти авторського права для навчання моделей ШІ з одночасним забезпеченням справедливої винагороди авторам і правовласникам. Такий підхід може стати альтернативою індивідуальному ліцензуванню, застосування якого ускладнюється масштабами навчальних масивів та необхідністю отримання дозволів від значної кількості правовласників.

Ще одним способом захистити твори від використання для навчання моделей ШІ є застосування спеціальних позначень і технічних засобів. Зокрема, автори можуть використовувати програму NO AI для маркування водяним знаком. Водночас власники вебресурсів можуть встановлювати технічні обмеження для автоматизованого збору їхнього контенту. Такі заходи застосовують і великі медіа: зокрема, The New York Times, BBC, CNN та Reuters обмежували доступ вебсканерів, що використовуються AI-компаніями для збору даних [9]. Подібні механізми дають правовласникам можливість більш чітко виразити незгоду з використанням їхніх творів для навчання моделей ШІ.

Зростання масштабів використання систем штучного інтелекту зумовило появу значної кількості судових спорів проти компаній-розробників у зв'язку з можливими порушеннями авторських прав під час навчання моделей ШІ. Водночас аналіз сучасної судової практики свідчить про відсутність усталеного підходу до вирішення таких спорів. Їх розгляд нерідко триває роками, а правовласники стикаються зі складнощами у доведенні факту та характеру використання своїх творів під час навчання моделей ШІ.

Показовою у цьому контексті є справа Andersen et al. v. Stability AI Ltd. et al. [10], ініційована у 2023 році групою художників проти розробників генеративних AI-сервісів. Позивачі стверджують, що їхні твори без дозволу були включені до масивів даних, використаних для навчання Stable Diffusion. Автори вимагали компенсації за завдані збитки через порушення авторських прав, правил DeviantArt, Закону про захист авторських прав у цифрову епоху (Digital Millennium Copyright Act), а також підтвердження того, що такі ситуації не повторюватимуться в майбутньому [11]. На початковому етапі суд відхилив значну частину заявлених вимог, звернувши увагу, зокрема, на недостатнє обґрунтування механізму копіювання творів. Після уточнення позову частина вимог щодо порушення авторського права продовжила розглядатися.

Важливим аспектом справи є розмежування можливого порушення авторського права на етапі навчання моделі та порушення, пов'язаного з результатами її роботи [10]. Відсутність істотної подібності між згенерованим зображенням і конкретним твором сама по собі не вирішує питання правомірності попереднього використання цього твору для навчання ШІ. При цьому оцінка такого використання значною мірою залежить від технічного механізму функціонування відповідної моделі.

Станом на серпень 2026 року остаточного рішення у справі не ухвалено. Водночас перебіг майже трирічного судового процесу дає змогу виокремити обставини, які мають значення для оцінки правомірності використання творів у межах доктрини fair use. Зокрема, враховуються спосіб отримання та використання твору, технічний механізм навчання моделі, а також можливість відтворення охоронюваних елементів твору у згенерованому результаті.

Одним із найбільш нових резонансних судових процесів у цій сфері є справа Encyclopaedia Britannica, Inc. та Merriam-Webster, Inc. v. OpenAI [12]. Позов було подано 13 березня 2026 року до Федерального окружного суду Південного округу Нью-Йорка, а станом на кінець серпня 2026 року розгляд справи триває.

Позивачі стверджують, що OpenAI без відповідного дозволу використовувала охоронюваний авторським правом контент для навчання своїх мовних моделей. Зокрема, Encyclopaedia Britannica володіє авторськими правами на колективні

твори, що охоплюють майже 100 000 онлайн-статей[13]. На думку позивачів, використання такого контенту для функціонування ChatGPT створює продукт, здатний частково замінювати першоджерело для користувача, що, своєю чергою, може призводити до зменшення відвідуваності ресурсів Britannica та Merriam-Webster і негативно впливати на їхні доходи.

Водночас претензії стосуються не лише етапу навчання моделей. У позові зазначається, що відповіді ChatGPT у певних випадках містять повні або часткові дослівні відтворення, перекази чи узагальнення охоронюваного контенту позивачів. Окремо оскаржується використання матеріалів Britannica у межах технології доповненої пошуком генерації (RAG), за допомогою якої система може звертатися до зовнішніх джерел під час формування відповіді користувачеві.

Справа становить особливий інтерес у контексті співвідношення ліцензування та доктрини fair use. OpenAI у спорах щодо використання загальнодоступного контенту для навчання моделей послідовно посилається на доктрину fair use. Водночас у цій справі суду належить оцінити заявлені позивачами порушення, тому питання щодо правомірності відповідного використання та можливості застосування fair use наразі залишається відкритим.

Слід зазначити, що цей позов є частиною ширшої хвилі судових спорів між онлайн-медіа та технологічними компаніями, які розробляють моделі штучного інтелекту. Подібні за правовими підставами та характером заявлених вимог позови були подані й іншими медіакомпаніями, що свідчить про формування окремого напряму судової практики щодо використання журналістського контенту для навчання та функціонування моделей ШІ.

Показовою для визначення меж правомірного використання творів під час навчання генеративного ШІ є справа GEMA v. Suno[14]. 31 липня 2026 року Земельний суд Мюнхена задовольнив більшість вимог німецької організації колективного управління GEMA до американського розробника генератора музики Suno. Предметом спору стало використання шести музичних творів із репертуару GEMA для навчання моделей відповідача та їх подальше відтворення у результатах роботи системи.

Одним із ключових висновків суду стало розмежування аналізу навчальних даних та їх «запам'ятовування» моделлю. Суд виходив із того, що у випадку закріплення охоронюваного змісту твору в параметрах моделі таким чином, що він надалі може бути відтворений у згенерованому результаті, може мати місце відтворення твору в розумінні авторського права.

Важливим є і підхід суду до винятку щодо глибинного аналізу текстів і даних. Суд визнав можливість його застосування до навчання генеративного ШІ як такого, проте відмежував необхідні для аналізу технічні копії від випадків, коли конкретний охоронюваний зміст зберігається у параметрах моделі та може бути надалі відтворений. На думку суду, таке «запам'ятовування» виходить за межі TDM-винятку. Окреме значення мало встановлення відсутності правомірного доступу до частини навчального контенту, отриманого шляхом обходу технічних обмежень на його завантаження.

Таким чином, рішення у справі GEMA v. Suno не заперечує можливості застосування винятків з авторського права до навчання моделей ШІ загалом, однак встановлює межу між аналізом охоронюваного контенту та фактичним збереженням його елементів у моделі з можливістю подальшого відтворення. Такий підхід має особливе значення для визначення співвідношення винятків з авторського права та ліцензування: чим ближчим результат функціонування моделі є до відтворення конкретного охоронюваного твору, тим складніше обґрунтувати відсутність необхідності отримання дозволу правовласника.

Висновки. Розвиток технологій штучного інтелекту є незворотним процесом, тому надмірне обмеження використання даних для навчання моделей ШІ може не лише стримувати інновації, а й створювати нерівні умови конкуренції між юрисдикціями з різним рівнем правових обмежень. Водночас технологічний розвиток не повинен відбуватися за рахунок необґрунтованого обмеження прав та економічних інтересів авторів і правовласників. Саме тому основним завданням правового регулювання має бути пошук балансу між цими інтересами.

На підставі проведеного дослідження до основних критеріїв оцінки правомірності використання об'єктів авторського права для навчання моделей ШІ доцільно віднести: правомірність доступу до твору; характер і обсяг його використання; наявність дозволу або застереження правовласника; можливість застосування передбачених законом винятків та обмежень авторського права; технічний механізм використання твору під час навчання; можливість відтворення охоронюваних елементів конкретного твору у результатах роботи

моделі; а також вплив такого використання на економічні інтереси правовласника, зокрема здатність згенерованого контенту конкурувати з оригінальним твором або замінювати його на відповідному ринку.

Аналіз судової практики свідчить, що формування критеріїв правомірності використання творів для навчання ШІ лише розпочинається. Такі судові процеси є тривалими та складними, зокрема через труднощі у встановленні факту використання конкретного твору та дослідженні технічних механізмів функціонування моделей.

Важливим напрямом удосконалення правового регулювання має залишатися підвищення прозорості навчання моделей ШІ. Правовласник повинен мати реальну можливість отримати інформацію про використання його творів та ефективно реалізувати передбачені законом способи захисту. У цьому контексті особливого значення набуває практична реалізація вимог AI Act щодо розкриття інформації про контент, використаний для навчання моделей загального призначення, а також удосконалення механізму opt-out.

Перспективним напрямом забезпечення балансу між інтересами правовласників та розвитком технологій ШІ є запровадження доступних механізмів ліцензування, зокрема колективних моделей, які дозволяли б використовувати значні масиви охоронюваного контенту без необхідності укладення окремого договору з кожним правовласником та водночас забезпечували б авторам справедливу винагороду. Хоча, звичайно, ліцензування не повинно бути обов'язковим для будь-якого використання творів у навчанні ШІ, але й fair use/TDM не повинні створювати універсальну презумпцію правомірності масового використання охоронюваного контенту. Поєднання прозорості навчальних даних, законодавчо визначених винятків, ефективного ліцензування та судового захисту може забезпечити належний баланс між охороною авторського права та подальшим розвитком технологій штучного інтелекту.

Список літератури:

1. 42 % дорослих і 70 % підлітків користуються ШІ в Україні: результати дослідження *Дія.Освіта*. Прес-офіс Міністерства. Міністерство цифрової трансформації України. URL:https://thedigital.gov.ua/news/education/42-doroslykh-i-70-pidlitkiv-korystuiutsia-shi-v-ukrayini-rezultaty-doslidzennia-diiaosvita (дата звернення: 24.08.2026)

2. How ChatGPT and our foundation models are developed. *OpenAI.* URL: https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed (accessed: 30.08.2026).

3. Директива Європейського Парламенту і Ради (ЄС) 2019/790 від 17 квітня 2019 року про авторське право і суміжні права на Єдиному цифровому ринку та про внесення змін до директив 96/9/ЄС та 2001/29/ЄС. *Відомості Верховної Ради України.* 2019. URL: https://zakon.rada.gov.ua/laws/show/984_022-19#Text (дата звернення: 31.08.2026).

4. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Official Journal of the European Union. 2024. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689 (accessed: 31.08.2026).

5. Researchers have trouble finding AI training data summaries. Euractiv. 2026. URL: https://www.euractiv.com/news/researchers-have-trouble-finding-ai-training-data-summaries (accessed: 21.08.2026).

6. UK creative industries face a clear and present danger from generative AI. *UK Parliament.* 2026. URL: https://committees.parliament.uk/committee/170/communications-and-digital-committee/news/212361/uk-creative-industries-face-a-clear-and-present-danger-from-generative-ai/ (accessed: 31.08.2026).

7. Петрів О. Штучний інтелект та авторське право. Центр демократії та верховенства права. 2023. URL:https://cedem.org.ua/analytics/shtuchnyi-intelekt-avtorske-pravo/ (дата звернення:22.08.2026).

8. Кульчій О.О. Інститут авторського права в умовах розвитку інструментів генеративного штучного інтелекту. *Науковий вісник Ужгородського національного університету.* Серія: Право. 2024. Вип. 86: частина 2. Ст. 143. DOI https://doi.org/10.24144/2307-3322.2024.86.2.22

9. Нановська В.The New York Times блокує вебсканер OpenAI (доповнено). *Медіамейкер*. 2023. URL: https://mediamaker.me/news/the-new-york-times-blokuye-vebskaner-openai/?utm_source=telegram&utm_medium=news&utm_campaign=vebskaner (дата звернення:10.03.2026).

10. Sarah Andersen, et al., v. Stability AI LTD., et al. United States District Court, Northern district of California. 2023. URL: https://fingfx.thomsonreuters.com/gfx/legaldocs/byprrngynpe/AI%20COPYRIGHT%20LAWSUIT%20mtdruling.pdf (accessed: 31.08.2026).

11. Петрів О., Спесивцева О. Автори проти штучного інтелекту: на чийому боці суд? Центр демократії та верховенства права. 2023. URL: https://cedem.org.ua/news/avtory-proty-shi-na-chyjomu-botsi-cud/?fbclid=IwAR1be8kM4gszuNsXl5l2hE85Om_u2Yz2ihnIWv0KeNzihcHo7fZzSNWHbP8 (дата звернення:10.08.2026).

12. Brittain B. Encyclopedia Britannica sues OpenAI over AI training. *Reuters*. 2026. URL: https://www.reuters.com/legal/litigation/encyclopedia-britannica-sues-openai-over-ai-training-2026-03-16/ (accessed: 31.08.2026).

13. Knobloch A. Copyright: Encyclopaedia Britannica sues OpenAI over AI training. *Heise Online*. 17 March 2026. URL: https://www.heise.de/en/news/Copyright-Encyclopaedia-Britannica-sues-OpenAI-over-AI-training-11213033.html (accessed: 31.08.2026).

14. Munich District Court Rules on AI-generated music: GEMA v Suno. *Bird & Bird.* 2026. URL: https://www.twobirds.com/en/insights/2026/germany/munich-district-court-rules-on-ai-generated-music-gema-v-suno (accessed: 31.08.2026).

Firman D. I. THE BALANCE BETWEEN THE FAIR USE DOCTRINE AND LICENSING IN ARTIFICIAL INTELLIGENCE MODEL TRAINING: LEGAL ISSUES SURROUNDING THE USE OF COPYRIGHTED WORKS

The article is devoted to a legal analysis of the use of copyright-protected works in the process of training artificial intelligence models and to identifying the criteria for determining the lawfulness of such use. It is established that the growing use of generative AI and the need for large volumes of data to train relevant models raise increasingly significant concerns regarding the protection of the rights and legitimate interests of authors and rightsholders.

The article analyses current approaches to regulating the use of copyright-protected works for AI training in the European Union and the United States, including the provisions of Regulation (EU) 2024/1689 (AI Act), text and data mining (TDM) exceptions, and the fair use doctrine. Particular attention is drawn to the cross-border nature of the use of training data, which complicates the determination of applicable law and jurisdiction in copyright infringement disputes.

It is established that the lawfulness of using copyright-protected works for AI model training cannot be determined solely by the fact that such works are publicly available on the Internet. Such an assessment should take into account the legal basis for access to the work, the nature and extent of its use, the existence of the rightsholder's authorisation or reservation of rights, the applicability of copyright exceptions and limitations, the model's ability to reproduce protected elements of the work, as well as the impact of such use on the rightsholder's economic interests. Attention is also given to the appropriateness of rightsholders using technical measures and specific markings to restrict the automated collection and use of their content for AI model training.

Special attention is paid to the opt-out mechanism and the challenges associated with its practical implementation, particularly the insufficient awareness of rightsholders regarding the possibility of reserving their rights against the use of their works for AI training. Current case law is examined, highlighting the difficulties in proving that specific works have been used in training datasets and in establishing the nature of such use.

Keywords: *copyright, artificial intelligence, AI model training, fair use, licensing, opt-out, digital environment.*

Дата першого надходження статті до видання: 31.08.2026
Дата прийняття статті до друку після рецензування: 14.09.2026
Дата публікації (оприлюднення) статті: 18.11.2026