

Створення, охорона, захист і комерціалізація об'єктів права інтелектуальної власності:
матеріали VIII Всеукраїнської науково-практичної конференції з міжнародною участю,
присвяченої Міжнародному дню інтелектуальної власності.
(м. Київ, 25 квітня 2025 р.): ел. збірник. Київ: КПІ ім. Ігоря Сікорського, 2025

УДК 347.77.01

ПРОБЛЕМАТИКА ЗАХИСТУ АВТОРСЬКИХ ПРАВ НА ОБ'ЄКТИ, ЩО БУЛИ ВИКОРИСТАНІ ПРИ МАШИННОМУ НАВЧАННІ

А.В. Тарасюк

аспірант НДІ інтелектуальної власності НАПрН України

ORCID ID 0009-0003-5015-8797

ISSUES OF COPYRIGHT PROTECTION FOR OBJECTS USED IN MACHINE LEARNING

A.V. Tarasiuk

**Ph.D. student of the Intellectual Property Scientific Research Institute of the
National Academy of Legal Sciences of Ukraine**

Анотація: У тезах розглянуто актуальну проблему захисту авторських прав у контексті використання об'єктів авторського права для навчання моделей штучного інтелекту. Автор аналізує технічні особливості процесу машинного навчання, що ускладнюють доведення факту порушення, а також висвітлює прогалини у чинному українському законодавстві та правозастосовній практиці.

Ключові слова: штучний інтелект, авторське право, машинне навчання, використання творів, правова охорона, інтелектуальна власність

Annotation: The paper addresses the pressing issue of copyright protection in the context of using copyrighted materials to train artificial intelligence models. The author analyzes the technical aspects of machine learning that complicate the identification and proof of infringement, as well as the gaps in current Ukrainian legislation and legal practice.

Keywords: artificial intelligence, copyright, machine learning, use of works, legal protection, intellectual property

Розвиток технологій штучного інтелекту означає нові виклики для охорони та захисту авторських прав. Однією з центральних тем, щодо якої точиться багато дискусій науковців та практиків є питання правомірності

використання об'єктів авторського права без згоди правовласників для навчання моделей штучного інтелекту.

Як можливу правову підставу такого використання часто визначають fair use та fair dealing (залежно від юрисдикції). Проте можливість використання вказаних концептів часто є сумнівною через реальну мету створення відповідних систем штучного інтелекту та ту роль, яку відведено відповідним об'єктам авторського права при їх створенні. Зокрема, компанія, що створює генератор реалістичних зображень може використати безліч фото з мережі інтернет для навчання своєї моделі штучного інтелекту, і, таким чином, створити, по суті програмне рішення, яке буде конкурентним для фотографів, чиї фотографії були використанні при навчанні відповідної моделі. Теж саме може стосуватись і авторів художніх книг.

Книги можуть бути «згодовані» певній моделі, а в результаті, система генеративного штучного інтелекту, в основі якої буде така модель – видаватиме користувачам цілі глави з відповідних книг, або ж створюватиме сіквели чи інші переробки / похідні твори, що значно зашкодить інтересам авторів, а може і змінить для них ринок в цілому.

Відповідно, цей новий тип використання об'єктів авторського права створив новий тип правовідносин щодо використання таких об'єктів в рамках машинного навчання, і, в рамках таких правовідносин, права авторів відповідних об'єктів можуть бути порушені і потребуватимуть захисту.

При цьому, реальний захист авторських прав таких авторів є складним та проблемним. По-перше, складним є виявлення факту використання відповідних об'єктів для тренування відповідних моделей. Якщо мова йде про системи генеративного штучного інтелекту, то таке виявлення може бути здійснене в рамках взаємодії з такою системою. Якщо ж справа стосується закритих систем, наприклад, систем, що використовують технології комп'ютерного зору, сказати точно, чи були використані певні об'єкти авторського права при їх навчанні доволі проблематично, якщо взагалі можливо.

По-друге, досудовий захист прав відповідних авторів є складним через відсутність сталої практики щодо алгоритмів дій учасників ринку. Використання процедури DMCA може бути одним з елементів захисту прав, але також є проблематичним

По-третє, відсутність сформованої судової практики щодо таких випадків означає складну прогнозованість рішень, що можуть бути винесені на підставі відповідних позовів та вплинути на рішучість авторів захищати свої права шляхом подачі до суду. В той же час, формування такої судової практики прямо зараз відбувається в США, адже саме там зібрані найбільші

гравці ринку бізнесу на з використанням штучного інтелекту і, скоріш за все, вирішення таких справ, матиме певний вплив на формування відповідної доктрини по всьому світу.

В зв'язку з цим, в цих тезах є необхідним дослідити деякі правові можливості захисту прав авторів, чиї роботи були використані неправомірно для навчання систем штучного інтелекту в українському правовому полі, з врахуванням трендів, що складаються в цій ніші в цілому.

Необхідним є розглянути правові та технологічні аспекти використання об'єктів авторського права при навчанні моделей штучного інтелекту

В рамках індустрії штучного інтелекту існує складність використання єдиного понятійного апарату, адже через технічну складність та значну динамічність розвитку технологій цієї групи деякі формулювання можуть по-різному розумітись учасниками правовідносин. Певну кодифікацію здійснило Міністерство цифрової трансформації, випустивши Словник термінів у сфері штучного інтелекту[1].

Окремо варто виділити термін «модель». Зокрема, визначається, що моделлю є «Програмне забезпечення або математичний вираз, що містить параметри, які можна налаштовувати і які визначають його поведінку. Ці параметри зазвичай налаштовуються за допомогою алгоритмів машинного навчання»[1].

Розуміння суті технічної операції, що має місце при машинному навчанні є ключовим для ефективного захисту прав авторів, які прагнуть припинити порушення їх авторського права в зв'язку з використанням об'єктів при машинному навчанні.

Отже, моделлю є певна багатошарова нейронна мережа з визначеними параметрами цієї мережі та певним комп'ютерним кодом, що дозволяє надсилати та отримувати запити до такої моделі. Цими параметрами, зокрема, є кількість шарів, а також ваги нейронів, індекс їх активації та зсуви. За своєю суттю в рамках моделі відбувається прогнозування певної відповіді, наприклад, чи є тварина на зображенні котом, або ж яким є наступний найбільш ймовірний символ / слово, яке має мати місце у відповіді на певне запитання.

Під час машинного навчання, навчальні матеріали, фактично пропускаються через шари нейронної мережі, змінюючи параметри, зокрема ваги нейронів, і таким чином, «допомагаючи» в майбутньому давати більш точні прогнози. Отже, якщо модель була «натренована» на книжках лише одного автора, то, відповідно, скоріш за все, система штучного інтелекту, в основі якої використовується ця модель буде надавати відповіді,

використовуючи цитати з цієї книги, або ж використовувати відповідний стиль і інші характерні речі при відповідях.

При цьому, самі книги автора не будуть певним каталогом, до якого здійснюються запити. Відповідна інформація про певні паттерни і особливості буде фактично вшита в параметри відповідної моделі.

Саме така технологічна структура моделі, і відповідно системи штучного інтелекту визначає складнощі при захисті авторських прав авторів, адже «дістати» певні об'єкти з системи є складним – потрібно або перетреновувати модель, або ж ставити певні блокери-запобіжники, які б не давали моделі видавати результати, що є такими, що ідентичні від творів, що були використанні при її навчанні.

Повертаючись до структури правовідносин, в якому має місце машинне навчання, варто зазначити, що об'єкти авторських прав можуть бути правомірно використані в машинному навчанні. Зокрема, відповідними підставами можуть бути ліцензійний договір, або ж використання доктрини чи відповідних положень fair use / fair dealing (але в цьому випадку, матимуть місце ризики). Також можливим є використання об'єктів, що є у суспільному надбанні. При цьому, доволі розповсюдженою практикою є пряма заборона (в публічній формі або в договорах використовувати об'єкти авторського права для тренування систем штучного інтелекту.

К. Зеров, зазначає, що «Згідно з українським законодавством дії з «тренування» систем штучного інтелекту можуть стосуватися права на відтворення твору — прямого чи опосередкованого виготовлення однієї або більше копій об'єкта авторського права та/або суміжних прав (або його частини) будь-яким способом та у будь-якій формі, у тому числі для тимчасового чи постійного зберігання в електронній (цифровій), оптичній або іншій формі, яку може зчитувати комп'ютер.»[2]

Відповідно, якщо авторське право автора, зокрема, на відтворення твору є порушеним, воно має підлягати захисту.

В зв'язку з цим необхідно більш детально розглянути правові аспекти захисту прав авторів, чиї твори були використані при навчанні моделей штучного інтелекту.

Міністерство цифрової трансформації України в кооперації з іншими організаціями випустило «Рекомендації щодо відповідального використання ШІ: питання права інтелектуальної власності. Міністерство цифрової трансформації України»[3].

В рамках цих Рекомендацій мають місце рекомендації авторам щодо захисту їх об'єктів авторського права в рамках їх неправомірного використання в машинному навчанні.

Зокрема, зазначається, що «Виходячи з положень статті 16 Цивільного кодексу України, частини 11 статті 33, статті 55 Закону України «Про авторське право і суміжні права», інших норм законодавства, правовласник, чий права інтелектуальної власності та/або права особливого роду порушені, може звернутися до суду за застосуванням ефективних способів захисту свого порушеного права».[3] та вказуються способи захисту авторського права, що визначені у статті 55 вищевказаного закону, а саме:

- заборону вчиняти дії, що порушують права правовласника;
- відшкодування моральної (немайнової) шкоди, завданої порушенням таких прав;
- відшкодування збитків, завданих порушенням таких прав;
- стягнення з порушника авторського права та/або суміжних прав доходу, отриманого внаслідок порушення;
- стягнення компенсації (разового грошового стягнення)[3].

За свою суттю, всі вказані способи є способами отримання фінансової компенсації, тоді як заборона вчиняти певні дії є конкретним приписом, що може бути застосований до відповідного порушника. Але от запитання – що це можуть бути за заборони? Наприклад – «не використовувати об'єкти з масиву, що додається на диску для тренування моделей X, Y.».

Цікавим цьому контексті видається спосіб захисту, що також передбачений статтею 55 – відновлення становища, яке існувало до порушення[3]. Таким відновленням може бути видалення «впливу» певних об'єктів на параметри відповідної моделі. При цьому, технічна реалізація цього припису є складною, якщо відповідач прагне і планує зберегти саму модель, як таку, при цьому, припинивши порушення прав автора.

В той же час, у вказаних рекомендаціях надаються доволі цікаві способи досудового захисту прав відповідних авторів, зокрема, подання скарги про порушення, медіація тощо[3].

В рамках як досудового, так і судового захисту прав авторів, чий твори були неправомірно використані при машинному навчанні важливим є збирання доказів. Одним з засобів доказування, що також зазначений, у вказаних вище Рекомендаціях, може бути використання промптів (запитів) до відповідних систем штучного інтелекту з метою отримання «відповідей», в яких би мало місце вираження певних творів, що певним чином, підтверджувало б їх використання при відповідному машинному навчанні, і, зокрема – можливої демонстрації виражень таких творів необмеженому колу осіб, які взаємодіють з відповідною системою штучного інтелекту. Зокрема, такий метод був використаний в рамках американського правосуддя в рамках

справи в Окружному суді США Південного округу Нью-Йорк, в рамках якої, позивачем є THE NEW YORK TIMES COMPANY, а відповідачами, inter alia, MICROSOFT CORPORATION, OPENAI, INC.[5].

Суть спору зводиться до претензій позивача щодо неправомірного використання відповідних авторських статей та матеріалів при машинному навчанні, в наслідок чого, їх «вираження» стало доступним користувачам Chat GPT, що означає, зокрема, збитки для Позивача. В рамках цієї справи, позивач додавав докази, як в рамках Chat GPT було отримано відповіді з відповідними цитатами, частинами тексту. Відповідач же зазначав про певний концепт «хакерського промпту, який не роблять звичайні користувачі».

Вимоги позивачів можуть варіюватись – від видалення конкретних об'єктів з дата-сету для навчання або ж блокування демонстрації користувачам відповідних систем штучного інтелекту «виражень» їх творів, до повного видалення всієї моделі, як такої, бо лише це зможе забезпечити ефективний захист їх прав.

В якості висновків необхідно зазначити, що захист прав авторів у випадку неправомірного використання їх творів при машинному навчанні з метою зміни параметрів відповідних моделей, що є частиною систем штучного інтелекту є проблематичним та складним в реалізації.

В той же час, використання передбачених законодавством способів захисту може призвести до задоволення вимог автора шляхом видалення відповідних об'єктів з датасетів або моделей (в тому числі шляхом перетренувань, дотренувань, або встановлення блокувань на видачу ідентичних об'єктам результатів), проте для успішної реалізації відповідних заходів з захисту, необхідним є розуміння технічних аспектів роботи ШІ систем.

Список використаної літератури

1. Словник термінів у сфері штучного інтелекту. Міністерство цифрової трансформації України. Посилання: [https://thedigital.gov.ua/storage/uploads/files/page/community/docs/%D0%A1%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA%20%D1%82%D0%B5%D1%80%D0%BC%D1%96%D0%BD%D1%96%D0%B2%20%D1%83%20%D1%81%D1%84%D0%B5%D1%80%D1%96%20%D0%A8%D0%86%20\(1\).pdf](https://thedigital.gov.ua/storage/uploads/files/page/community/docs/%D0%A1%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA%20%D1%82%D0%B5%D1%80%D0%BC%D1%96%D0%BD%D1%96%D0%B2%20%D1%83%20%D1%81%D1%84%D0%B5%D1%80%D1%96%20%D0%A8%D0%86%20(1).pdf)
2. Зеров К. Чи мріють системи генеративного штучного інтелекту про електричних овець? Поняття та умови охороноздатності об'єктів, згенерованих системами генеративного штучного інтелекту в Україні.

Створення, охорона, захист і комерціалізація об'єктів права інтелектуальної власності:
матеріали VIII Всеукраїнської науково-практичної конференції з міжнародною участю,
присвяченої Міжнародному дню інтелектуальної власності.
(м. Київ, 25 квітня 2025 р.): ел. збірник. Київ: КПІ ім. Ігоря Сікорського, 2025

Теорія і практика інтелектуальної власності. 2023. No 4–5. С. 90–100.

Посилання:

https://drive.google.com/file/d/1w1rtL6ADlCM_u8FRWFSrKyjRRg2KoISt/view

3. Рекомендації щодо відповідального використання ІІІ: питання права інтелектуальної власності. Міністерство цифрової трансформації України. Посилання: chrome-<https://thedigital.gov.ua/news/vikoristovuyte-shtuchniy-intelekt-bez-porushennya-prava-intelektualnoi-vlasnosti-rekomendatsii-dlya-rozrobnikov-koristuvachiv-i-pravovlasnikov-kontentu>
4. Про авторське право і суміжні права: Закон України від 1 грудня 2022 року. No 2811-IX. URL: <https://zakon.rada.gov.ua/laws/show/2811-20#Text>
5. Матеріали справи No 1:23-cv-11195. Окружний суд США Південного округу Нью-Йорк.

Посилання:

<https://www.courtlistener.com/docket/68117049/parties/the-new-york-times-company-v-microsoft-corporation/>