

ДО ЗАПРОВАДЖЕННЯ ПРАВА НА ВИКОРИСТАННЯ ТВОРІВ У МАШИННОМУ НАВЧАННІ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

Іван ВАЩИНЕЦЬ

*кандидат юридичних наук, доцент,
керівник Науково-консультативного центру (м. Ужгород)
Київського університету права Національної академії наук України
м. Ужгород, Україна*

Бурхливий розвиток систем штучного інтелекту (ШІ) актуалізував питання правової кваліфікації використання охоронюваних авторським правом творів у процесі машинного навчання. У доктрині та практиці висловлюються діаметрально протилежні підходи: від заперечення охоплення таких дій змістом виключних прав до вимоги отримувати дозвіл правовласників і сплачувати їм винагороду. У судах різних юрисдикцій розглядаються спори з приводу такого використання, проте прийняті рішення наразі не дають підстав для якихось обґрунтованих узагальнень. Для вітчизняного авторського права це питання є так само актуальним з огляду на рекодифікацію цивільного права України та адаптацію національного законодавства до права ЄС.

Для з'ясування того, чи машинне навчання охоплюється змістом виключних прав, необхідно попередньо встановити, які саме дії з об'єктами авторського права здійснюються під час такого навчання. Відповідні дослідження на замовлення компетентних органів проведені останнім часом в США, Європі та під егідою Організації економічного співробітництва та розвитку. Так, ОЕСР у травні 2024 року видала роботу «Штучний інтелект, дані та конкуренція», де розглянула новітні розробки у сфері ШІ та їхній життєвий цикл, у т.ч. й навчання базових моделей^[1]. У травні 2025 року під егідою Офісу з авторського права США опублікований звіт, присвячений питанням використання об'єктів авторського права для розвитку генеративних систем ШІ, який є черговою, третьою частиною у низці досліджень щодо впливу ШІ на авторське право^[2]. У тому ж місяці вийшло дослідження Офісу з інтелектуальної власності Європейського Союзу «Розвиток

^[1] OECD (2024). *Artificial intelligence, data and competition*. OECD Artificial Intelligence Papers, No. 18. OECD Publishing, Paris. 62 p. <https://doi.org/10.1787/e7e88884-en>.

^[2] Perlmutter, S. & Vidal, K. (2025). *Copyright and Artificial Intelligence Part 3: Generative AI Training*. United States Copyright Office. United States of America. 113 p. <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>

генеративного штучного інтелекту з позицій авторського права»^[3], а місяцем пізніше – дослідження, проведене на замовлення Департаменту політики з питань юстиції, громадянських свобод та інституційних справ Європейського парламенту «Генеративний ШІ та авторське право – навчання, створення, регулювання»^[4].

Указані дослідження дають достатні підстави вважати, що дії, які вчиняються під час машинного навчання систем ШІ, охоплюються змістом виключних прав на твір. Відповідно, за загальним правилом таке використання має здійснюватися з дозволу правовласника та супроводжуватися виплатою авторської винагороди.

Так, машинне навчання розглядається як використання статистичних алгоритмів для виявлення закономірностей у даних і, відповідно, для автономного набуття знань (самостійного навчання)^[5]. Отримані дані система ШІ обробляє шляхом т.зв. «токенізації», тобто поділу тексту на окремі слова, їх частини, цифри та знаки пунктуації, які стають одиницями даних (токенами), та обчислює мільярди вагових коефіцієнтів шляхом статистичного визначення того, який токен з найбільшою ймовірністю буде слідувати за відповідним попереднім токеном^[6]. Подібним чином опрацьовуються візуальні дані.

Описані дії спрямовані на забезпечення належної якості результатів, які продукує система ШІ. Як зазначається в одному із згаданих досліджень, генеративні системи не просто виділяють статистичні кореляції з даних; вони опрацьовують і моделюють експресивні патерни, закладені в об'єктах, охоронюваних авторським правом, з метою генерування результатів – таких як текст або зображення, – які можуть нагадувати контент, створений людиною^[7]. Тобто під час машинного навчання здійснюються аналіз, синтез, відтворення та трансформація даних. При цьому у результаті цього процесу вагові коефіцієнти у системах ШІ можуть містити копії охоронюваних творів з навчальних даних. У такому випадку подальше копіювання цих коефіцієнтів навіть стороною, не залученою у процес навчання, може становити *prima facie*

^[3] EUIPO (2025). *The Development of Generative Artificial Intelligence from a Copyright Perspective*. 436 p. <https://www.euipo.europa.eu/es/publications/genai-from-a-copyright-perspective-2025>

^[4] European Parliament: Directorate-General for Citizens' Rights, Justice and Institutional Affairs. *Generative AI and copyright – Training, creation, regulation*. European Parliament, 2025. 175 p. <https://data.europa.eu/doi/10.2861/9120512>

^[5] OECD (2024). *Artificial intelligence, data and competition*. OECD Artificial Intelligence Papers, No. 18. OECD Publishing, Paris. P. 13. <https://doi.org/10.1787/e7e88884-en>

^[6] EUIPO (2025). *The Development of Generative Artificial Intelligence from a Copyright Perspective*. P. 147, 149. <https://www.euipo.europa.eu/es/publications/genai-from-a-copyright-perspective-2025>

^[7] European Parliament: Directorate-General for Citizens' Rights, Justice and Institutional Affairs. *Generative AI and copyright – Training, creation, regulation*. European Parliament, 2025. P. 78. <https://data.europa.eu/doi/10.2861/9120512>

порушення авторських прав^[8]. Це ж стосується і випадків, коли вихідні матеріали, згенеровані ШІ, є тотожними чи значно подібними до охоронюваних творів^[9]. Крім того, у процесі машинного навчання можливе порушення права суб'єкта авторських прав на переробку створеного ним твору^[10]. Йдеться про випадки, коли вихідні матеріали, згенеровані ШІ, є переробкою використаних для навчання об'єктів авторського права.

Отже, у процесі машинного навчання систем ШІ здійснюється ряд операцій з навчальними даними, у т.ч. й об'єктами авторського права, які відповідають способам використання цих об'єктів, що входять до змісту виключних прав інтелектуальної власності на твір.

У певних випадках є підстави вважати, що таке використання творів може підпадати під встановлені винятки й обмеження з авторського права. Саме посилання на визначені винятки або на доктрину добросовісного використання (fair use) часто становить основу заперечень відповідачів у спорах про порушення авторського права системами ШІ. Водночас сам по собі факт використання творів саме у процесі машинного навчання не дає підстав для такої кваліфікації. Як вказує Офіс з авторського права США, вирішення питання про те, чи таке використання є добросовісним, залежатиме від того, які саме твори використовувалися, з якого джерела, з якою метою та за наявності яких механізмів контролю результатів їх використання у вихідних даних моделі. Водночас комерційне використання величезних масивів об'єктів авторського права для створення експресивного контенту, який конкурує з ними на наявних ринках, особливо якщо це досягається шляхом неправомірного (незаконного) доступу, виходить за межі усталених підходів до застосування доктрини добросовісного використання^[11].

Подібний підхід існує в праві ЄС із глибинним аналізом тексту та даних (text and data mining, TDM), який є винятком з авторських прав. Хоча TDM та генеративний ШІ можуть спиратися на певні спільні технічні методи, їхні функції та результати принципово відрізняються. TDM спрямований на вилучення аналітичних висновків із даних; генеративний ШІ стосується алгоритмічного синтезу результатів, що імітують або відтворюють експресивні (виражальні) патерни, наявні в уже створеному контенті^[12]. Як наслідок, і серед науковців-правників, і

^[8] Perlmutter, S. & Vidal, K. (2025). *Copyright and Artificial Intelligence Part 3: Generative AI Training*. United States Copyright Office. United States of America. P. 28. <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>

^[9] Ibid., p. 31.

^[10] Ibid., p. 29.

^[11] Ibid., p. 107.

^[12] European Parliament: Directorate-General for Citizens' Rights, Justice and Institutional Affairs. *Generative AI and copyright – Training, creation, regulation*. European Parliament, 2025. P. 40. <https://data.europa.eu/doi/10.2861/9120512>

в середовищі фахівців з технологій обґрунтовується підхід, згідно з яким навчання генеративних моделей ШІ за своєю суттю не є різновидом TDM – ані з технічної, ані з юридичної точки зору^[13].

Однією з принципових причин таких активних дискусій з досліджуваного питання є те, що під час машинного навчання доступ до об'єктів авторського права та дії щодо цих об'єктів вчиняються без участі людини. Це дало можливість стверджувати, що за відсутності доступу і використання людиною таких об'єктів не може бути порушень авторського права. Один із американських дослідників зазначав, що «роботизовані читачі» не можуть порушувати авторське право, і ми так само не дозволяємо роботам бути авторами, оскільки авторське право – для людей^[14]. Проте з такою позицією важко погодитися через низку причин, найголовнішою з яких є той факт, що виведення за межі дії авторського права використання творів у машинному навчанні завдає шкоди не тільки суб'єктам авторського права, які не отримують за таке використання винагороди, але може зашкодити розробникам систем ШІ, які навряд чи працюватимуть належним чином без безперервного надходження створених людиною результатів творчої діяльності, з яких вони можуть навчатися^[15].

Наявність численних спірних питань кваліфікації використання творів у процесі машинного навчання зумовила появу в доктрині пропозицій щодо запровадження окремого виключного права. Наявні конструкції авторського права – як система виключних прав, так і винятків та обмежень – виявляються недостатніми для забезпечення належного балансу між інтересами правовласників та розробників систем ШІ. У зв'язку з цим видається обґрунтованою позиція тих дослідників, які пропонують запровадити окреме виключне право на використання творів під час машинного навчання систем ШІ, так зване «learnright»^[16]. З урахуванням складності досліджуваної проблеми слід визнати доцільним створення певної системи законодавчих положень (learnright law), покликаній визначити нове виключне право щодо надання дозволу автоматизованим системам навчатися на об'єктах, які охороняються авторським правом. Такий підхід відповідав би як правовим засадам, так і моральним цінностям, на яких ґрунтується авторське право. У той же час запровадження нового виключного права повинно супроводжуватися уточненням системи винятків і обмежень з існуючих майнових прав інтелектуальної власності на твір, зокрема щодо глибинного аналізу тексту та даних.

^[13] Ibid., p. 42.

^[14] James Grimmelman (2016). Copyright for Literate Robots. *Iowa Law Review*. 2016. Volume 101. P. 673, 681.

^[15] Frank Pasquale, Thomas W. Malone, and Andrew Ting (2025). Copyright, Learnright, and Fair Use: Rethinking Compensation for AI Model Training. *Northwestern Journal of Technology and Intellectual Property*. Volume 23. Issue 1. P. 208.

^[16] Ibid., p. 213-221.