



**Глотов Сергій**  
**кандидат юридичних наук,**  
**науковий співробітник**  
**НДІІВ НАПрН України, адвокат**  
**Харків, Україна**  
**<https://orcid.org/0000-0002-6496-4162>**  
**e-mail: [glotov.sergiy@gmail.com](mailto:glotov.sergiy@gmail.com)**

УДК 347.7

<https://doi.org/10.69724/2786-8834-2025-4-1-72-89>

## АВТОРСЬКЕ ПРАВО V. ШТУЧНИЙ ІНТЕЛЕКТ: ЧИ Є МАШИННЕ НАВЧАННЯ ВИКОРИСТАННЯМ?

**Глотов Сергій. Авторське право v. штучний інтелект: чи є машинне навчання використанням?**

### **Анотація**

Статтю присвячено питанню машинного навчання штучного інтелекту в аспекті оцінювання такого, насамперед технічного, процесу з погляду авторського права, а саме: чи слід вважати, що в цьому випадку здійснюється використання творів. Ураховуючи, що таке машинне навчання звичайно проводиться за рахунок охоронюваних творів, а також неодноразовість ставлення до цього питання в суспільстві, що, з одного боку, виявляється в гучних позовах до розробників штучного інтелекту на кшталт *The New York Times v. OpenAI & Microsoft Inc.*, а з іншого – у різного роду соціальних ініціативах, скерованих на сприяння розвитку інструментів штучного інтелекту та поглиблення розуміння громадськістю переваг штучного інтелекту, розглянута проблема є край важливою та цікавою. В статті приділено увагу останнім тенденціям у сфері законодавчих ініціатив з акцентом на право Європейського Союзу. Зокрема, дано стислу характеристику нещодавно ухваленому Регламенту (ЄС) 2024/1689 від 13 червня 2024 р., який ще називають Законом ЄС про штучний інтелект. Зауважується, що в разі навчання штучного інтелекту за рахунок творів, які перейшли в суспільне надбання, вихідні результати штучного інтелекту відповідатимуть рівню знань сторічної давнини, тобто будуть неактуальними. Крім того, старі твори можуть бути придатні до використання в процесі машинного навчання лише в разі їх оцифрування, що можна зробити тільки вручну. Це гальмуватиме процес розвитку новітніх технологій в усіх сферах. Обґрунтовано, що те, що відбувається під час машинного навчання штучного інтелекту, являє собою не що інше, як «читання»: особливу форму користування твором, право автора на яку не поширюється. Встановлено, що користування твором з метою машинного навчання підпадає під дію обмеження щодо глибинного аналізу тексту та даних, регламентоване статтею 4 Директиви 2019/790/ЄС. Доведено, що культурна творчість немислима без використання попередніх досягнень інших авторів. А тому, оскільки вихідний результат штучного інтелекту не є правооб'єктом, в контексті машинного навчання штучного інтелекту тут немає жодних дій, що є значущими з погляду авторського права.

**Ключові слова:** *штучний інтелект, авторське право, машинне навчання, вихідний результат ШІ, технологія.*

### Аналіз останніх досліджень і публікацій

Окрім аспектів і проблемні питання, пов'язані із штучним інтелектом, охоплюються науковими інтересами вчених різних країн, таких як О. Сімсон, О. Баранов, К. Зеров, Ю. Жорнокуй, К. Дюрантай, А. Левендовські, В. Кретцг, Т. Дорніс, М. Бойрскенс та ін. Так чи інакше в наукових дослідженнях останніх часів порушувались питання, пов'язані з машинним навчанням. Однак у всьому різноманітті наукових пошуків відсутня усталеність підходів в оцінюванні цього явища.

### Постановка проблеми

Системи штучного інтелекту (далі – ШІ) щодня значно впливають на суспільство в цілому та повсякденне життя людини зокрема. Водночас, попри очевидні переваги цього зручного допоміжного інструменту, він, як виявилось, може створювати ризики та завдавати шкоди суспільним інтересам і основним правам, які захищені законом. Така шкода може бути матеріальною чи нематеріальною, включаючи фізичну, психологічну, соціальну чи економічну шкоду.

Як відомо, ключовою характеристикою систем ШІ є їх здатність робити висновки. Ця здатність робити висновок стосується процесу отримання вихідних даних, таких як прогнози, вміст, рекомендації чи рішення, які можуть впливати на фізичне та віртуальне середовище, а також здатності систем штучного інтелекту отримувати моделі чи алгоритми, або те й інше, із вхідних даних, або дані. Техніки, які дають змогу зробити висновок під час побудови системи ШІ, включають підходи до машинного навчання, які навчаються на основі даних, як досягти певних цілей, і підходи, засновані на логіці та знаннях, які роблять висновок із закодованих знань або символічного представлення завдання, яке потрібно вирішити. Здатність системи ШІ робити висновки виходить за рамки базового опрацювання даних, надаючи можливість навчатися, міркувати або моделювати. Термін «машинний» стосується того факту, що системи ШІ працюють на машинах.

Для того, щоб система ШІ у вихідному результаті (*output content*) видавала максимально релевантний результат, його необхідно навчати за допомогою сучасних високоякісних творчих та інтелектуальних здобутків – те, що називається «*input content*», переважна частина якого охороняється авторським правом. Чим слабше навчальний матеріал, тим більшої кількості помилок буде припускатися система ШІ. Ось чому суспільство має бути зацікавлене в тому, щоб ШІ навчався на сучасних високоякісних навчальних матеріалах.

Однак через те, що більшість систем ШІ навчається із застосуванням охоронюваних авторським правом здобутків без отримання на це згоди правовласників, автори та інші суб'єкти авторського права схильні вважати, що навчання ШІ за допомогою належних їм творів становить порушення виключного права на відтворення, отже, використання є значущим з погляду авторського права. Станом на сьогодні немає усталеної правової позиції щодо того, чи є навчання систем ШІ з використанням охоронюваних творів таким, що не порушує авторських прав. Отже, з цього питання існує правова невизначеність.

**Метою** цієї статті є демонстрація правової невизначеності в оцінюванні такого технічного процесу, як машинне навчання ШІ, через очевидний конфлікт з авторським правом у зв'язку із необхідністю застосування в процесі навчання охоронюваних творів. **Завдання**, яке ставиться при цьому, полягає в тому, щоб відповісти на питання: чи є застосування охоронюваних творів використанням, що є значущим з погляду авторського права.

### Викладення основного матеріалу

Коли мова заходить на тему права і технологій, мабуть, перше, що спадає на думку, є «ШІ» – термін, щодня дедалі частіше використовується в літературі, науці та мистецтві. Здається, навіть, що цей термін відомий багатьом, хоча при цьому складається стійке враження, що суспільство тут явно ще не перебуває на піку свого ажіотажу.

За своєю семантикою термін «ШІ» настільки ж туманний, наскільки і суперечливий. Це явище за своєю суттю насправді є й «штучним», й «інтелектуальним» одночасно. У сфері авторського права ШІ звичайно належить до інтелекту, який «імітується» комп'ютерами, в якому «розумна» поведінка людини моделюється за допомогою (як правило, простих) алгоритмів.

А втім, що взагалі являє собою інтелект у живих істот? За вікіпедією, це здатність розуміти зв'язки між ідеями, встановлювати зв'язки та вирішувати проблеми, використовуючи доступну інформацію. Це передбачає здатність не лише здобувати знання, а й ефективно застосовувати їх у різноманітних контекстах. Інтелект не обмежується лише академічною чи аналітичною діяльністю, а й поширюється на емоційне розуміння, соціальні взаємодії, «здоровий глузд» та практичні навички вирішення проблем.

Комп'ютери, дійсно, блискучі в тому, щоб швидко обробляти великі обсяги даних. І хоча «відсутність чіткого визначення ШІ не заважає оцінювати інтелектуальність на інтуїтивному рівні», ШІ у певному сенсі можна порівняти з губкою, яка вбирає величезну кількість інформації, але не завжди розуміє її сенс. Його основна сила полягає у здатності аналізувати величезні масиви даних і знаходити в них закономірності. Саме це пояснює, чому розвиток машинного навчання останнім десятиріччям сприяв значним досягненням у сфері прогнозування тексту, автоматичного перекладу, розпізнавання зображень, голосу та інших напрямках, де опрацювання великих обсягів інформації дає змогу ефективно розв'язувати складні завдання. Проте станом на сьогодні немає остаточної відповіді на питання, чи має значення машинне навчання з погляду авторського права.

Машинне навчання – це здатність машин отримувати знання з даних. Проте навіть «глибинне навчання» («*Deep Learning*») не може вишколити ШІ на засвоєння мистецтва створення оригінальних внутрішніх логічних взаємозв'язків, а лише взмозі будувати гіпотези, спираючись на велику кількість ілюстративного матеріалу. Тому навряд чи ШІ можна вважати інтелектом у людському розумінні. Адже машини не є по-справжньому автономними у своєму «мисленні» та «ухваленні рішень».

Подібно до того, як губка поглинає воду без розрізнення її якості, ШІ аналізує дані, знаходить закономірності та використовує їх для надання вихідного результату (*output content*), хоча самостійно не розуміє значення інформації так, як людина. Цим, власне, й пояснюється, чому нейромережі можуть генерувати логічно правильні тексти чи передбачати результати, але при цьому іноді роблять помилки через відсутність справжнього усвідомлення контексту.

Думається, зрештою, Йоханес Масінг (Johannes Masing) має рацію: термін «штучний інтелект» не повинен ставати надто самостійним, оскільки це просто форма особливо розвиненої обробки даних і програмування<sup>1</sup>.

У будь-якому випадку описане фізичне явище з огляду на свою здатність породжувати юридично значущі наслідки, причому в різних сферах права, зокрема цивільному, кримінальному, інтелектуальному, трудовому та міжнародному праві, потребує свого нормативно-правового регулювання<sup>2</sup>.

Звертаючись до законодавства ЄС, що, зважаючи на панівні в Україні євроінтеграційні процеси, є для нас більш затребуваним, логічним було б припустити, що питання, пов'язані зі ШІ, врегульовані Директивою 2019/790/ЄС «Про авторське право і суміжні права на Єдиному цифровому ринку та про внесення змін до директив 96/9/ЄС та 2001/29/ЄС» від 17 квітня 2019 р. (Директива 2019/790 або DSM-Директива)<sup>3</sup>. Однак у DSM-Директиві, якою врегульовано так звані TDM-бар'єри<sup>4</sup> на рівні ЄС, поняття «ШІ» не згадується. Проте це не має викликати

<sup>1</sup> Masing J. Balance wahren. Zeitschrift für Rechtspolitik (ZRP). 2020. S. 194.

<sup>2</sup> Наприклад, постають такі запитання: хто є автором твору, створеного за допомогою ШІ? Хто відповідатиме за шкоду, отриману людиною внаслідок неправильного діагнозу, поставленого ШІ? Хто відповідатиме за кримінальне правопорушення, скоєне ШІ, наприклад, за Deepfake-маніпуляції, які спричиняють наклеп або шахрайство? Чи має бути взагалі законодавчо встановлене зобов'язання повідомляти третіх осіб про використання ШІ? І так далі.

<sup>3</sup> Директива 2019/790/ЄС «Про авторське право і суміжні права на Єдиному цифровому ринку та про внесення змін до директив 96/9/ЄС та 2001/29/ЄС» від 17 квітня 2019 р. URL: [https://zakon.rada.gov.ua/laws/show/984\\_022-19#Text](https://zakon.rada.gov.ua/laws/show/984_022-19#Text) (дата звернення: 20.01.2025).

<sup>4</sup> Йдеться про обмеження щодо інтелектуального аналізу текстів і даних (TDM – Text and Data Mining), які стосуються відтворення, пов'язаного з навчанням. Частина 2 ст. 2 DSM-Директиви визначає TDM як метод автоматизованого аналізу тексту та даних у цифровій формі з метою вилучення інформації, включаючи (але не обмежуючись) закономірності, тенденції та кореляції. Стаття 3 DSM-Директиви передбачає обмеження на використання TDM для наукових досліджень, а ст. 4 DSM-Директиви – для комерційного використання.

Варто зазначити, що деякі автори припускають, що навчання генеративного ШІ підпадає під бар'єри TDM (див. тут, наприклад: Theodoros Chiou. Copyright lessons on Machine Learning: what impact on algorithmic art? 2019. 10 JIPITEC, pp. 398–411, p. 409, marginal 22; Griffiths J., Synodinou Tatiana, Xalabarder Raquel. Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of Directive (EU) 2019/790 on Copyright in the Digital Single Market. (2023) 72 GRUR Int. 22, p. 25; Senftleben Martin. Generative AI and Author Remuneration, (2023) 54 IIC 1535, p. 1542 at fn. 33; Vesala Juha. Developing Artificial Intelligence-Based Content Creation: Are EU Copyright and Antitrust Law Fit for Purpose?, (2023) 54 IIC 351, p. 355; Hamann Hanjo. Artificial Intelligence and the Law of Machine-Readability: A Review of Human-to-Machine Communication Protocols and their (In) Compatibility with Article 4(3) of the Copyright DSM Directive, 15 (2024) JIPITEC 102 para 1, p. 105). З іншого боку, слід зауважити, що в силу тодішньої неактуальності цього питання здатність генеративного

подив, оскільки генеративний ШІ став відомий широкому загалу лише у 2022 р., спочатку через генератори зображень, а потім – у листопаді 2022 р. – через ChatGPT.

21 квітня 2021 р. Комісією ЄС було надано свій проєкт європейського закону про ШІ<sup>5</sup>, де офіційно було вперше сформульовано, що поняття «система ШІ» розуміється як:

- 1) програмне забезпечення;
- 2) те, що має на меті низку *цілей*, поставлених людьми;
- 3) те, що може давати результати, такі як контент, прогнози, рекомендації чи рішення, які впливають на середовище, з яким вони взаємодіють;
- 4) те, що розроблено із використанням однієї чи кількох із наведених нижче *технік і концепцій*<sup>6</sup>:

– *концепції машинного навчання*, з контрольованим, неконтрольованим і підкріплювальним навчанням з використанням широкого спектру методів, включаючи глибоке навчання;

– *логіка та концепції, засновані на знаннях*, включаючи подання знань, індуктивне (логічне) програмування, бази знань, машини логічного висновку та дедукції, (символічні) міркування та експертні системи;

– *статистичні підходи, байєсівське оцінювання, методи пошуку та оптимізації*.

Пізніше з ухваленням Регламенту (ЄС) 2024/1689 від 13 червня 2024 р. це розуміння, що встановлює гармонізовані правила щодо ШІ та вносить зміни до Регламентів (ЄС) № 300/2008, (ЄС) № 167/2013, (ЄС) № 168 /2013, (ЄС) 2018/858, (ЄС) 2018/1139 і (ЄС) 2019/2144 і Директиви 2014/90/ЄС, (ЄС) 2016/797 і (ЄС) 2020/1828 (далі – Регламент (ЄС) 2024/1689)<sup>7</sup>, який складається з 180 пунктів преамбули до нього та 113 статей, трансформувалося в таке:

*«Система штучного інтелекту» означає машинну систему, яка розроблена для роботи з різними рівнями автономності та може виявляти адаптивність після розгортання, і яка, для явних або неявних цілей, робить висновки на основі вхідних*

---

штучного інтелекту зберігати, відтворювати та перетворювати захищені авторським правом твори будь-якого роду все ж ще була за межами уяви законодавця того часу. А тому є точка зору, що бар'єри TDM не можна сприймати як приписи для генеративного ШІ.

<sup>5</sup> Пропозиція щодо регламенту Європейського парламенту та Ради, що встановлює гармонізовані правила щодо штучного інтелекту (Закон ЄС про штучний інтелект) із внесенням змін до деяких правових актів Союзу (див.: URL:[https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC\\_1&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF) (дата звернення: 28.01.2025)).

<sup>6</sup> Додаток 1 до проєкту Закону ЄС про штучний інтелект під назвою «Техніки і концепції ШІ». URL: [https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC\\_2&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_2&format=PDF) (дата звернення: 28.01.2025).

<sup>7</sup> URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689> (дата звернення: 21.01.2025). Регламент (ЄС) 2024/1689 набуває чинності в усіх 27 країнах-членах ЄС з 1 серпня 2024 р., а виконання більшості його положень почнеться з 2 серпня 2026 р.

даних, які вона отримує, як генерувати результати, такі як прогнози, вміст, рекомендації або рішення, які можуть впливати на фізичне або віртуальне середовище (ст. 3(1) Регламенту (ЄС) 2024/1689)<sup>8</sup>.

Взагалі Регламент (ЄС) 2024/1689, який ще називають Законом про ШІ, є вельми цікавим і дійсно важливим документом, адже він фактично став першою всеосяжною горизонтальною правовою основою для регулювання систем ШІ в ЄС. З огляду на те, що цей законодавчий акт є результатом довготривалих переговорів, спрямованих на створення узгодженої правової бази «для розробки, розміщення на ринку, введення в експлуатацію та використання систем штучного інтелекту»<sup>9</sup> в ЄС, а також оскільки цей документ є відносно новим, дамо йому тут невеличкий огляд.

Насамперед слід зазначити, що Регламент використовує ризико-орієнтований підхід до регулювання всього життєвого циклу різних типів систем ШІ. За недотримання його приписів може бути накладено фінансовий штраф у розмірі до 35 мільйонів євро або (максимально) 7 відсотків світового річного обороту, залежно від того, яка сума більша.

Наведене вище визначення поняття «система ШІ», по суті, має на меті відрізнити ШІ від більш простих програмних систем. Регламентом встановлюються зобов'язання для постачальників, дилерів<sup>10</sup>, імпортерів, дистриб'юторів і виробників систем ШІ з прив'язкою до ринку ЄС. Наприклад, Закон ЄС про ШІ застосовується до: (i) постачальників, які розміщують на ринку ЄС або вводять в експлуатацію системи ШІ або розміщують на ринку ЄС моделі ШІ загального призначення<sup>11</sup>; (ii) розробники систем ШІ, які мають місцезнаходження/знаходяться в ЄС; і (iii) постачальників і розробників систем ШІ в третіх країнах, якщо результати, створені системою ШІ, використовуються в ЄС (ст. 2(1) Регламенту (ЄС) 2024/1689). Закон ЄС про ШІ також перераховує певні винятки щодо його матеріальної сфери застосування (наприклад, Регламент не застосовується до систем ШІ з відкритим кодом, якщо вони не заборонені або класифікуються як системи ШІ високого ризику або системи ШІ, що використовуються виключно з метою наукових досліджень та розробок) (пп. 3, 4, 6, 8, 10 та 12 ст. 2 Регламенту (ЄС) 2024/1689).

Закон ЄС про ШІ забороняє певні практики ШІ в ЄС, які він вважає шкідливими, образливими та такими, що суперечать цінностям ЄС. Заборонені методи

<sup>8</sup> Таке визначення повністю узгоджується із визначенням цього терміна, наданим Організацією економічного співробітництва та розвитку (англ. The Organisation for Economic Co-operation and Development (OECD)) у своїх рекомендаціях: OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, доступних за посиланням: URL: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (дата звернення: 21.01.2025).

<sup>9</sup> Див. п. 1 Преамбули до Регламенту (ЄС) 2024/1689.

<sup>10</sup> deploy, з англ. розгортати.

<sup>11</sup> Так звані «моделі GPAI», щодо яких Регламент містить окрему главу для їх класифікації та регулювання.

ШІ включають застосування підсвідомих методів ШІ за межами свідомості людини або цілеспрямованих маніпулятивних чи оманливих методів з метою чи наслідком суттєвого спотворення людської поведінки. Однак він передбачає кілька винятків із цього правила для цілей правоохоронних органів, пов'язаних із використанням віддаленої біометричної ідентифікації «в реальному часі» в загальнодоступних місцях (ст. 5(2) Регламенту (ЄС) 2024/1689).

Із метою запровадження пропорційного та ефективного набору правил для систем ШІ Закон ЄС про ШІ встановлює підхід до регулювання, що ґрунтується на оцінці ризиків і класифікує системи ШІ на основі інтенсивності та масштабу ризиків, які може створити кожна система ШІ. «Системи ШІ високого ризику», тобто системи, що становлять «високий» ризик, належать до двох категорій: (i) системи ШІ, що використовуються як компонент безпеки продукту (або іншим чином підпадають під дію законодавства ЄС про гармонізацію здоров'я та безпеки); та (ii) системи ШІ, розгорнуті у восьми конкретних сферах, включаючи (серед іншого) освіту, працевлаштування, доступ до основних державних і приватних послуг, правоохоронну діяльність, міграцію та відправлення правосуддя (ст. 6(1) – (2) і Додаток III «Системи штучного інтелекту високого ризику, про які йдеться в статті 6(2)» Регламенту (ЄС) 2024/1689).

Регламентом регулюються питання стосовно так званих «відвертих фейків» (*deepfake*), які визначаються як «згенеровані штучним інтелектом або оброблені зображення, аудіо- чи відеоконтент, який нагадує існуючих осіб, об'єкти, місця, сутності чи події та може здаватися людині автентичним або правдивим» (ст. 3(60) Регламенту (ЄС) 2024/1689). Відповідно до Закону ЄС про ШІ діпфейкери, які використовують системи штучного інтелекту для створення діпфейків, зобов'язані чітко повідомити, що вміст було штучно створено або оброблено шляхом позначення результату ШІ як такого та розкриття його штучного походження (якщо використання не дозволено законом для виявлення, попередження, розслідування та переслідування кримінального правопорушення). Якщо вміст є частиною явно художнього твору, зобов'язання щодо прозорості обмежуються розкриттям існування такого створеного або обробленого вмісту так, щоб не перешкоджати показу чи насолоді від твору (ст. 50(4) Регламенту (ЄС) 2024/1689).

Беремо на себе сміливість стверджувати, що у зв'язку з величезною соціальною напруженістю, яка є добре помітною в публічній площині, найбільше проблем при використанні ШІ виникає з авторським правом. Серед питань, які здебільшого хвилюють суспільство, ключовими звичайно є питання про те, хто несе відповідальність, якщо використовуються результати ШІ, схожі на інші твори? Або ж: чи захищений результат, отриманий за допомогою ШІ, авторським правом? Або ж: чи допустимо навчання ШІ за допомогою творів без згоди їх авторів?

Очевидно, що найбільш болючим є питання, пов'язане із навчанням ШІ.

Як відомо, ШІ потребує повноцінного тренування. Передусім для цього потрібна велика база даних. Якість тренування ШІ залежить від якості матеріалу, на яко-

му він навчається. Тільки в разі, якщо вхідні дані містили достатню кількість текстів, зображень і музичних творів відповідної якості, можна в подальшому розраховувати на отримання якісного контенту у вигляді текстів, зображень, музичних творів і т. ін. Недарма в ІТ-середовищі та світі комп'ютерних наук існує незаперечний принцип: «garbage in, garbage out»<sup>12,13</sup>.

Цілком природно, що тексти та зображення, які здебільшого використовуються для навчання, зазвичай охороняються авторським правом. Наприклад, мовна модель штучного інтелекту LLaMA (Large Language Model Meta AI) від Meta AI навчалася, серед іншого, зі змістом книг із проєкту Гутенберга<sup>14</sup>. Мабуть, тому більшість розробників не розкривають, які саме навчальні дані були використані. А деякі набори навчальних даних, такі як BookCorpus, не мають задокументованого переліку творів, які вони містять<sup>15</sup>.

Вимоги до охороноздатності результату інтелектуальної творчої діяльності людини в цілому не можна назвати суворими, тим більше враховуючи те, що охорона авторським правом по суті надається на фідучіарній основі за заявою творця. Тобто потрібно зовсім небагато, щоб людська творчість стала об'єктом охорони авторським правом. Отже, можна припустити, що більшість контенту, який ШІ використовує для навчання, перебуває під охороною авторським правом<sup>16</sup>. Більше того, використання контенту, захищеного авторським правом, є необхідним, тому що в протилежному випадку це неминуче погіршить вихідні результати ШІ. Якби ШІ навчався виключно на творах, що знаходяться у суспільному надбанні, то його формулювання могли б здаватися такими, ніби вони походять із XIX – початку XX ст.<sup>17</sup> ШІ використовував би застарілі стандарти та, ймовірно, мав би ще більшу необ'єктивність у своїх відповідях, ніж це іноді відбувається зараз. Загалом, ШІ буде більш схильним до помилок, якщо матиме навчальні матеріали гіршої якості. Тому існує суспільний інтерес, щоб навчати ШІ на сучасному та якісному матеріалі. Адже

<sup>12</sup> У перекладі з англ. «сміття на вході – сміття на виході».

<sup>13</sup> Katharina de la Durantaye. Garbage In, Garbage Out. Regulating Generative AI Through Copyright Law. ZUM 10/2023, pp. 645–660, p. 647. URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4572952](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4572952) (last accessed: 27.01.2025).

<sup>14</sup> Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothee Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, Guillaume Lample. LLaMA: Open and Efficient Foundation Language Models, 2023. pp. 1–27. URL: <https://arxiv.org/pdf/2302.13971> (last accessed: 23.01.2025)

<sup>15</sup> Bandy Jack, Vincent Nicholas. Addressing «Documentation Debt» in Machine Learning Research: A Retrospective Datasheet for BookCorpus, 2021. [arXiv:2105.05241]. URL: <https://arxiv.org/abs/2105.05241> (last accessed: 30.01.2025)

<sup>16</sup> Hofmann Franz. Zehn Thesen zu Künstlicher Intelligenz (KI) und Urheberrecht. Zeitschrift für Wettbewerb in Recht und Praxis, 2024. Heft 1, S. 11–18, Rn. (Punkt) 10

<sup>17</sup> Kraetzig Viktoria. KI-Kunst als schöpferische Zerstörung, Neue Juristische Wochenschrift (NJW) – 2024, S. 697–702, Rn. (Punkt) 8



навіть якщо до розвитку ШІ ставитися критично, зупинити його вже неможливо – принаймні не засобами авторського права.

Сучасні мовні моделі ШІ звичайно використовують так звану трансформерну архітектуру, де «перетворення нейронних послідовностей відбувається за структурою кодера-декодера»<sup>18</sup>: кодер відображає вхідну послідовність символічних зображень  $(x_1, \dots, x_n)$  у послідовність неперервних зображень  $z = (z_1, \dots, z_n)$ ; далі за заданим  $z$  декодер генерує вихідну послідовність  $(y_1, \dots, y_m)$  символів по одному елементу за раз. На кожному кроці модель є авторегресійною<sup>19</sup>, використовуючи попередньо згенеровані символи як додаткові вхідні дані для генерації наступних. Для таких трансформерних моделей уже неодноразово було доведено, що вони постійно зберігають тренувальні дані<sup>20</sup>. Дослідники з Університету Берклі (Berkeley University) змогли довести, що GPT-4 знає дослівно велику кількість романів і може доповнювати імена в текстових уривках<sup>21</sup>. В опублікованому списку книг, які GPT-4 знає особливо добре, міститься багато творів, захищених авторським правом, зокрема твори Дж. К. Роулінг та Дж. Р. Р. Толкіна<sup>22</sup>.

На тлі вимушеного, але несанкціонованого використання охоронюваних об'єктів авторського права не дивно, що розробники великих моделей ШІ, які нині мають значні фінансові ресурси, що є для правовласників вирішальним, стикаються з численними судовими процесами щодо порушення авторських прав. Визначальною на міжнародному рівні є справа, подана The New York Times проти OpenAI та Microsoft у федеральний суд Нью-Йорка<sup>23</sup>. Ще в листопаді це судове відомство відхилило позов Raw Story Media, Inc. та Alternet Media, Inc. проти OpenAI<sup>24</sup>. Однак ухвалені рішення стосувалося передусім процесуального питання щодо допустимості розгляду справи. Відповідно до американського законодавства позивач повинен не лише довести порушення федерального права, а й обґрунтувати конкретну індивідуальну шкоду, завдану таким порушенням. У випадку із Raw Story Media

<sup>18</sup> Vaswani Ashish, Shazeer Noam and other. Attention Is All You Need. arXiv:1706.03762v7 [cs.CL] 2 Aug 2023. pp.1–15. URL: <https://arxiv.org/pdf/1706.03762> (last accessed: 29.01.2025).

<sup>19</sup> Graves Alex. Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850, 2013. URL: <https://arxiv.org/abs/1308.0850> (last accessed: 29.01.2025).

<sup>20</sup> Matthew Jagielski, Om Thakkar, Florian Tramer, Daphne Ippolito, Katherine Lee, Nicholas Carlini, Eric Wallace, Shuang Song, Abhradeep Thakurta, Nicolas Papernot, Chiyuan Zhang. Measuring Forgetting of Memorized Training Examples. arXiv:2207.00099v2 [cs.LG] 9 May 2023. URL: <https://arxiv.org/pdf/2207.00099> (last accessed: 30.01.2025).

<sup>21</sup> Kent K. Chang, Mackenzie Cramer, Sandeep Soni, David Bamman. Speak, Memory: An Archaeology of Books Known to ChatGPT/GPT-4. arXiv:2305.00118v2 [cs.CL] 20 Oct 2023. URL: <https://arxiv.org/pdf/2305.00118> (last accessed: 30.01.2025).

<sup>22</sup> Список доступний тут: URL: <https://github.com/bamman-group/gpt4-books> (last accessed: 30.01.2025).

<sup>23</sup> The New York Times Co. v. Microsoft Corp., et al., No. 23-CV-11195 (S. D. N. Y.). URL: <https://casetext.com/case/the-ny-times-co-v-microsoft-corp-3> (дата звернення: 05.01.2025).

<sup>24</sup> Raw Story Media, Inc. v. Openai Inc., Civil Action 1:24-cv-01514-CM (S. D. N. Y. Sep. 25, 2024). URL: <https://casetext.com/case/raw-story-media-inc-v-openai-inc-3> (last accessed: 05.01.2025).

не йшлося про безпосереднє копіювання контенту, а лише про видалення інформації про права (зокрема, імені автора, назви твору та застережень щодо прав). Суддя аргументував своє рішення тим, що при величезному обсязі обробленої інформації видалення таких даних у процесі генерації відповідей великими мовними моделями завдає настільки незначної індивідуальної шкоди конкретному право-власнику, що вимога компенсації у грошовій формі є необґрунтованою. Крім того, ймовірність, що конкретний уривок твору буде відтворено у відповідях на запити користувачів (особливо у великій кількості випадків), вважається вкрай низькою.

У Німеччині значну увагу привернув судовий спір між LAION e. V. та фотографом Робертом Кнешке (Robert Kneschke) про порушення авторських прав шляхом аналізу фотографій при створенні набору даних для навчання ШІ, який розглядав Земельний суд Гамбурга (LG Hamburg)<sup>25</sup>. Суд першої інстанції підтвердив широке тлумачення положень Закону про авторське право щодо текстового та дата-майнінгу (Text- and Datamining). Крім того, в листопаді 2024 р. організація колективного управління правами GEMA подала позов до Земельного суду Мюнхена I (LG Munchen I) проти OpenAI, L. L. C. та OpenAI Ireland Ltd., звинувативши компанії в нелегальному використанні текстів пісень у чат-боті.<sup>26</sup> Стосовно цього кейсу є цікавим одне з пояснень GEMA щодо підстави позову: «Численні відомі німецькі музиканти, зокрема Крістіна Бах (Атемлос), Рольф Цуковські, Райнхард Мей, Інґа Хампе, Томмі Екарт, Ульф Зоммер і Пітер Плейт, а також їхні музичні видавці підтримують позов GEMA. Доведено, що чат-бот використав тексти їх пісень. Компенсації за це вони не отримали»<sup>27</sup>.

Реагуючи на судові тяганини та ризики, із цим пов'язані, OpenAI уклала кілька ліцензійних угод з постачальниками контенту, зокрема з Axel Springer<sup>28</sup>, Wall Street Journal<sup>29</sup>, Financial Times<sup>30</sup>, Associated Press<sup>31</sup> та Bertelsmann<sup>32</sup>. Як правило, відповідні угоди також передбачають право контент-провайдерів використовувати техно-

<sup>25</sup> Rechtsprechung LG Hamburg, 27.09.2024–310 O 227/23.

URL: <https://betriebs-berater.ruw.de/wirtschaftsrecht/urteile/Zur-Urheberrechtsverletzung-durch-Analyse-von-Lichtbildern-beim-Aufbau-eines-KI-Trainingsdatensatzes-46743> (last accessed: 02.01.2025)

<sup>26</sup> URL: <https://www.gema.de/de/w/gema-erhebt-klage-gegen-openai> (last accessed: 21.01.2025)

<sup>27</sup> Ibid. «Zahlreiche namhafte deutsche Musikschaefende, darunter Kristina Bach (Atemlos), Rolf Zuckowski, Reinhard Mey, Inga Humpe, Tommi Eckart, Ulf Sommer und Peter Plate sowie deren Musikverlage unterstützen die Klage der GEMA. Ihre Songtexte wurden nachweislich durch den Chatbot verwertet. Eine Vergütung haben sie dafür nicht erhalten».

<sup>28</sup> URL: <https://openai.com/index/axel-springer-partnership> (last accessed: 24.01.2025).

<sup>29</sup> URL: <https://openai.com/index/news-corp-and-openai-sign-landmark-multi-year-global-partnership> (last accessed: 24.01.2025).

<sup>30</sup> URL: <https://openai.com/index/content-partnership-with-financial-times>(last accessed: 24.01.2025).

<sup>31</sup> URL: [www.axios.com/2023/07/13/ap-openai-news-sharing-tech-deal](http://www.axios.com/2023/07/13/ap-openai-news-sharing-tech-deal) (last accessed: 24.01.2025).

<sup>32</sup> URL: <https://www.bertelsmann.com/news-and-media/news/bertelsmann-and-openai-agree-on-strategic-collaboration.jsp?atn=2862170&abp=2862170,2862222> (last accessed: 24.01.2025).

логії ШІ. Однак лише фінансово сильні комерційні провайдери можуть дозволити собі такі підстраховні угоди. А це неминуче призведе до певної монополізації ресурсу, залишаючи інших дрібних або некомерційних контент-провайдерів, включаючи, до речі, й наукову спільноту, в аутсайдерах.

Очевидно, що без маркування технічно неможливо ідентифікувати вміст, який не захищений законом, чи тому, що він не є результатом власної інтелектуальної творчої праці, чи тому, що не відповідає критеріям оригінальності, чи тому, що строк їх охорони закінчився. Не можна оминати увагою й те, що ті твори, які стали суспільним надбанням, а це, як правило, твори віком не менше, аніж сто років, перед їх використанням для навчання ШІ необхідно оцифрувати вручну. А такою честою ушановано лише невелику кількість людських творчих здобутків.

Тому основним питанням, на якому слід зосередитися, є те, чи вчиняються під час навчання дії, що мають значення з погляду авторського права і чи можуть вони бути виправдані, тобто включені до випадків вільного використання (без виплати винагороди). Відповідям на ці питання було присвячено експертний висновок Тіма Дорніса та Себастьяна Штобера (Tim W. Dornis, Sebastian Stober)<sup>33</sup>, підготовлений ними за дорученням Ініціативи «Авторське право» (Authors' Rights Initiative). Цей висновок являє собою міждисциплінарний аналіз, де дослідження технологічної бази (с. 23–66) було написано Себастьяном Штобером, а Тім Дорніс відповідав за частини дослідження, присвячені правовим питанням (с. 67–192).

Висновок Дорніса/Штобера сколихнув не тільки професійний світ, а й широку громадськість: згідно з ним, створення моделей ШІ, а отже, і тим більше пропозиція систем ШІ на їх основі та їх використання, є незаконним, якщо використовується контент, захищений авторським правом. Це означає, що при навчанні ШІ, який генерує текст, музику та зображення, існує не лише загроза позову про відшкодування збитків, а й, можливо, навіть кримінальної відповідальності, принаймні на території окремих країн ЄС (наприклад, Німеччини).

Професійна громадськість відреагувала на цей експертний висновок низкою критичних публікацій, в одній з яких проф. Міхаель Бойрскенс (Prof. Dr. Michael Beurskens) із певною часткою іронії поставив запитання: чи не варто в такому випадку розробникам ШІ емігрувати до США чи Китайської Народної Республіки?<sup>34</sup> А наприкінці проф. Бойрскенс на прикладі Німеччини робить головний висновок: якщо матиме місце повна заборона на використання всіх моделей ШІ, навчених за допомогою охоронюваних творів, це значно відкине німецьку економіку в купі з освітою та наукою<sup>35</sup>.

<sup>33</sup> Tim W. Dornis, Sebastian Stober. Urheberrecht und Training generativer KI-Modelle: Technologische und juristische Grundlagen. Series: Recht und Digitalisierung | Digitization and the Law – NOMOS Verlag : Band 19, 1. Auflage 2024. 217 S. URL: [https://www.nomos-elibrary.de/10.5771/9783748949558.pdf?download\\_full\\_pdf=1](https://www.nomos-elibrary.de/10.5771/9783748949558.pdf?download_full_pdf=1) (last accessed: 14.01.2025).

<sup>34</sup> Michael Beurskens Training generativer KI nur auf Lizenzgrundlage? Eine Analyse des rechtlichen Gutachtens von Dornis/Stober. Zeitschrift für Recht Digital (RDİ). 2025 – Heft 1, S.1–7.

<sup>35</sup> Ibid., S.7.

З іншого боку, обурення авторів та правовласників охоронюваного контенту є в цілому зрозумілими. Прямуючи за логікою, яка притаманна економіці права, власники авторських прав стверджують, що за рахунок охоронюваних творів та чи інша модель ШІ отримує додану вартість, до якої зацікавлені особи за підсумком не мають жодного стосунку. А враховуючи те, що ШІ принаймні частково здатен замінити творчу роботу людини, оригінальні твори, як і самі творці, будуть нікому не потрібні. І якщо це так, якщо автори мають залишитися без роботи завдяки власним творам, вони мають бути принаймні фінансово залучені до такого типу використання<sup>36</sup>, адже такий спосіб використання твору – «годування» ШІ з метою машинного навчання – має всі ознаки незалежного економічного значення, що виходить за межі вільного використання.

Проте мусимо не погодитися з наведеними вище тезами з таких міркувань.

На нашу думку, те, що відбувається під час машинного навчання ШІ, становить не що інше, як «читання». За своєю суттю читання – це особлива форма *користування* твором, право автора на яку не поширюється. Автор також не може забронити особі споживати твір<sup>37,38</sup>. І таке споживання не має диференціюватися залежно від того, в якому просторі воно відбувається: аналоговому чи цифровому. Тобто цей принцип має однаково працювати і там, і там<sup>39</sup>. А за словами Катаріни де ла Дюрантай (Katharina de la Durantaye), при такому споживанні комерційно використовується лише інформація, що міститься в творі, але не особистий інтелектуальний творчий здобуток<sup>40</sup>.

Далі, з точки зору європейського авторського права збереження контенту, який є «їжею» для ШІ, в перетвореному стані є відтворенням у значенні ст. 2 Директиви 2001/29/ЄС «Про гармонізацію окремих аспектів авторського права і суміжних прав в інформаційному суспільстві» від 22 травня 2001 р.<sup>41</sup> Однак у цьому випадку цілком може бути задіяне обмеження щодо глибинного аналізу тексту та даних – бар'єр TDM, встановлений ст. 4 DSM-Директиви.

<sup>36</sup> Max Dregelies. KI-Training unter dem AI Act. GRUR 2024, S.1487

<sup>37</sup> Franz Hofmann. Grundsatz der Technikneutralität im Urheberrecht? Zugleich Gedanken zu einem more technological approach. Jahrgang 8 (2016) / Heft 4, S.492.  
URL: [https://viewer.content-select.com/pdf/viewer?id=1&id\\_type=doi&identifiers=10.1628%2f186723716X14846459813187&signature=eab02b31cf0e661ff38baf3fb6d2ccc63d2efb7e&session=00000000-0000-0000-0000-000000000000](https://viewer.content-select.com/pdf/viewer?id=1&id_type=doi&identifiers=10.1628%2f186723716X14846459813187&signature=eab02b31cf0e661ff38baf3fb6d2ccc63d2efb7e&session=00000000-0000-0000-0000-000000000000) (last accessed: 20.01.2025).

<sup>38</sup> Зрозуміло, що від початку особа, яка навчає ШІ, повинна мати законний доступ до твору.

<sup>39</sup> Хоча той же Франц Хофманн передбачає, що із диджиталізацією настане «кінець свободи в насолоді від твору» («Ende der Freiheit des Werkgenusses»). Ibid., S.498

<sup>40</sup> Katharina de la Durantaye. Nutzung urheberrechtlich geschützter Inhalte zum Training generativer künstlicher Intelligenz – ein Lagebericht. Zeitschrift für das gesamte Medienrecht (AfP), 1/2024, S. 9–22, Rn. (Punkt) 22.

<sup>41</sup> URL: [https://zakon.rada.gov.ua/laws/show/984\\_005-01#Text](https://zakon.rada.gov.ua/laws/show/984_005-01#Text) (last accessed: 07.01.2025).

Ну й нарешті, якщо порівняти ШІ з електронними особами – нагадаємо, що проєкт такої концепції був свого часу наданий Комітетом Європейського парламенту з правових питань<sup>42</sup>, згідно з яким роботів можна класифікувати як «електронних осіб» і навіть наділити їх правами та обов'язками, – то машинне навчання ШІ можна подати як освітній процес, аналогічний освітньому процесу людини. Тоді з певною часткою умовності і процес підготовки ШІ на запит можна порівняти з творчим процесом людини, з тією лише – суттєвою – відмінністю, що результат ШІ не є правооб'єктом<sup>43</sup>. Якщо це так, то до оцінки машинного навчання слід підходити як до навчання людського, так само як і до «народжуваних» результатів.

Дозволимо собі зауважити, що класичне християнське вчення про творіння *ex nihilo* у нашому житті не працює. Жодне творіння не виникає з нічого. Причиною цього є те, що культурна творчість немислима без використання попередніх досягнень інших авторів. Натхнення митця має джерело, чи то в природі, чи то у світі мрій митця, чи то в античній статуй, чи то у світі байок давно забутих часів. Робота з чужими творами і черпання з них натхнення є частиною суті інтелектуальної та творчої діяльності й може бути продемонстрована на незліченних прикладах. Художня і духовна творчість людини завжди базувалася на тому, що її оточувало. Людина завжди спирається на знання своїх предків, але вона не обмежується джерелами минулого, а використовує також джерела сьогодення і майбутнього. Багато творів літератури та образотворчого мистецтва створюються завдяки

<sup>42</sup> Доповідь з рекомендаціями до Комісії з норм цивільного права у галузі робототехніки 27.01.2017–2015/2103(INL). URL: [https://www.europarl.europa.eu/doceo/document/A-8-2017-0005\\_DE.html?redirect](https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_DE.html?redirect) (дата звернення: 15.01.2025).

Ю. Жорноку у своєму дослідженні в цілому обстоє позицію наділення ШІ правосуб'єктністю, хоча й вважає схожу новелу такою, що поки що є недоцільною через складність та проблематичність цього питання для вирішення (див.: Жорноку Ю. Порівняльний аналіз правосуб'єктності юридичної особи та штучного інтелекту. *Цивілістична платформа*. 2024. № 3. С. 98–112).

<sup>43</sup> Останнім часом проблема щодо можливості захисту результатів діяльності, які не були створені в рамках класичного творчого процесу людиною, набула особливої актуальності. Виникають питання, чи можуть і якою мірою результати, отримані з використанням ШІ, підлягати охороні авторським правом, а також: чи може такий ШІ вважатися автором (творцем). Термін «ШІ» часто розуміється як «розумна поведінка людини, змодельована алгоритмами» (Dorothee Thum in: Wandtke/Bullinger. *Praxiskommentar Urheberrecht: UrhR. C. H. BECK Verlag*. 6., neu bearbeitete und erweiterte Auflage – 2022. XXXVII, 3091 S., S.138, Rn.16). Схожа ситуація склалася і в англійських країнах, де усталеного визначення терміна «ШІ» ще не існує (Marketa Trimble, *Artificial Intelligence and Human Intelligence, GRUR International*, Vol. 72, Issue 1, January 2023, P. 1–2). Питання охороноздатності виникає насамперед у правових системах, які менше переймаються всебічним захистом автора, але більшою мірою захистом інвестицій, як, наприклад, англо-американська система авторського права. Як відомо, в авторському праві творцем може бути лише людина, а отже, ШІ не може бути автором. Питання про те, чи є результати ШІ творами взагалі, в будь-якому випадку є передчасним, оскільки авторське право виникає лише щодо особистих інтелектуальних творчих результатів. Згідно з переважаною думкою, «чисті» результати ШІ не є такими творами, оскільки в них повністю відсутня інтелектуально-людська комунікація з продуктом, яка є передтечею для будь-якого законодавства про авторське право (Dorothee Thum in: Wandtke/Bullinger. *Praxiskommentar Urheberrecht: UrhR. C. H. BECK Verlag*. 6., neu bearbeitete und erweiterte Auflage – 2022. XXXVII, 3091 S., S.138, Rn.18). Тому результати, створені за допомогою ШІ, не підлягають авторсько-правовій охороні.

тому, що їх автори отримують натхнення (інспіруються) більш ранніми творіннями; у музиці варіації або фантазії на чужу тему є поширеною формою композиції. Причому використовуються не тільки вільні від охорони авторським правом, а й захищені твори. Використання чужих творів у цьому контексті виправдане тим, що автор, на чий твір він «претендує», так само спирався або міг спиратися на чужі творіння. Плоди людської праці є результатом постійного обміну<sup>44</sup>.

Так, є всі підстави для того, щоб машинне навчання ШІ залишити поза межами уваги авторського права. У світлі цього, до речі, будь-які звинувачення від одного розробника ШІ до іншого через користування тими самими матеріалами для машинного навчання<sup>45</sup>, навіть шляхом їх копіювання<sup>46</sup>, виглядають не дуже слушними.

Насправді більш актуальними та важливими є питання щодо можливого порушення авторських прав унаслідок вихідних результатів ШІ. Особливо це стосується випадків, коли вихідний результат ШІ імітує стиль творця (той же пастиш, наприклад) або запозичує частини захищених творів. Дослідження показують, що ШІ цілком може генерувати результати, які частково містять захищений авторським правом матеріал<sup>47</sup>. Проте, на наше переконання, це має вирішуватися на машинному рівні: розробники систем ШІ повинні інтегрувати у системи ШІ захисні механізми, які запобігатимуть створенню правопорушного вихідного результату, наскільки це практично можливо. Іншими словами, тут «відповідь на машину — сама машина»<sup>48</sup>.

Як у людському світі, так і, звичайно, у випадку із ШІ, художній стиль не охороняється авторським правом так само, як абстрактна ідея. Проблеми з авторським правом виникають, якщо результат містить частини вже наявного твору, захищеного авторським правом. У цьому випадку можливе порушення права автора на від-

<sup>44</sup> Зрозуміло, що таке використання має бути обмежене. Це досягається тим, що має бути створений незалежний новий твір і що це має бути вільне використання, за якого характеристики використаного твору зникають з огляду на індивідуальність нового твору.

<sup>45</sup> Microsoft перевіряє, чи не отримувала група, пов'язана з DeepSeek (нещодавно заявлений інструмент ШІ китайського походження), дані OpenAI неналежним чином. URL: <https://www.reuters.com/technology/microsoft-probing-if-deepseek-linked-group-improperly-obtained-openai-data-2025-01-29/> (дата звернення: 31.01.2025).

<sup>46</sup> Тут, звісно, не можна оминати того, що збірка творів, за допомогою яких було навчено ШІ та які збережені всередині нього, начебто може претендувати на охорону як база даних. Проте це є неможливим у правовому сенсі через функціонально-технічні та математико-статистичні характеристики (CJEU. Case C-604/10 Football Dataco v. Yahoo! ECLI:EU:C:2012:115).

<sup>47</sup> Див., наприклад: Getty Images (US), Inc. v. Stability AI, Inc. (1:23-cv-00135). У цій судовій справі Getty Images подала позов, звинувачуючи Stability AI у порушенні прав на понад 12 мільйонів фотографій, пов'язаних із ними підписів і метаданих під час створення та пропозиції Stable Diffusion і DreamStudio. Ця справа також включає звинувачення в порушенні прав на торговельну марку, пов'язані зі здатністю «обвинуваченої» технології відтворювати водяні знаки Getty Images у генеративних результатах ШІ. URL: <https://www.courtlistener.com/docket/66788385/getty-images-us-inc-v-stability-ai-inc/>

<sup>48</sup> Charles Clark. The Future of Copyright in a Digital Environment. Kluwer Law International, 1996, pp. 139–146.

творення або права на адаптацію твору. Однак стосовно сучасних моделей ШІ, на наш погляд, це має місце лише в тих випадках, коли ШІ, будучи навченим за допомогою відповідного твору, обробляє запит, який стосується саме цього твору. Якщо, з іншого боку, ШІ створює продукт, який суто випадково схожий на інший твір без навчання ШІ на цьому творі, то це буде так званім подвійним створенням. А це не є порушенням авторських прав<sup>49</sup>.

Більш поширеним є випадок, коли вихідний результат ШІ нагадуватиме вже наявні твори, інформацію про які було включено в нейронну мережу як частину навчання. Водночас тут буде важливим, наскільки продукт ШІ схожий на інший твір. Іншими словами, чи можуть ті компоненти вихідного твору, які відображають особистість автора, бути розпізнаними у вихідному результаті ШІ.

За будь-яких обставин, ми не маємо забувати, що всі без винятку вихідні результати ШІ не є правооб'єктами. А це означає, що за підсумком відповідальність за порушення авторських прав має покладатися виключно на особу, яка стала в подальшому використовувати цей машинний результат, видаючи його за твір.

## Висновки

Економічні переваги систем ШІ незаперечні – вони заощаджують робочий час, можуть компенсувати фізичні та розумові вади та сприяють тому, щоб зробити інформацію більш доступною. Однак із сучасним станом технологій розробка необхідних систем ШІ залежить від великих обсягів даних. Проте якщо класифікувати машинне навчання як акт використання твору, який не охоплюється жодними винятками та обмеженнями й потребує згоди, власники прав на такі твори стануть вартівниками для доступу до необхідної інформації. Тим більше, що окремі автори можуть заборонити глибинний аналіз тексту та даних по своїх творах, користуючись приписом п. 3 ст. 4 DRM-Директиви. Розробники ШІ практично позбавлені можливості заздалегідь отримати всі необхідні відповідні права – принаймні, за економічно обґрунтовану ціну<sup>50</sup>. Проте, ми переконані, що навіть, для прикладу, найтиражніший роман за всіх часів незалежності України Василя Шкляра «Чорний ворон» не може бути недоторканим настільки, щоб не ставати доступним для машинного навчання. Крім того, так чи інакше будь-якого одного твору для машинного навчання замало. ШІ його навіть не помітить. У світі цього суспільства опиняється на вістрях наступної соціальної дилеми: тоді як кожен автор вважає свій внесок значним, цінність може бути створена шляхом компіляції величезної кількості окремих елементів. А це означає, що цінність кожного окремого елементу є мінімальною.

З огляду на викладене вище, обстоюємо позицію, що вільна доступність, зокрема і захищених творів, для машинного навчання систем ШІ, враховуючи нікчемність

<sup>49</sup> Ulrich Loewenheim in: G. Schricker. Urheberrecht: Kommentar. C. H. BECK Verlag – 3 Auflage, 2006, S. 72, Rn. 42.

<sup>50</sup> Kraetzig V. KI-Kunst als schöpferische Zerstörung. Neue Juristische Wochenschrift (NJW). 2024, S. 697–702, Rn. (Punkt) 9.

таких дій із точки зору авторського права, є не тільки юридично обґрунтованою, а й служить інтересам світової громадськості.

## REFERENCES

- Annex 1 to the Draft EU AI Act titled «Techniques and Concepts of AI». URL: [https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC\\_2&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_2&format=PDF) (accessed: January 28, 2025) [in English]
- Bandy, J., & Vincent, N. (2021). Addressing «documentation debt» in machine learning research: A retrospective datasheet for BookCorpus. Retrieved from: <https://arxiv.org/abs/2105.05241> [in English]
- Beurskens, M. (2025). Training generativer KI nur auf Lizenzgrundlage? Eine Analyse des rechtlichen Gutachtens von Dornis/Stober [Training generative AI only on a licensing basis? An analysis of the legal opinion by Dornis/Stober], *Zeitschrift für Recht Digital (RD*i*)*, 1, 1–7 [in German]
- Chang, K. K., Cramer, M., Soni, S., & Bamman, D. (2023). Speak, memory: An archaeology of books known to ChatGPT/GPT-4. *arXiv:2305.00118v2* [cs.CL] [in English]
- Chiou, T. (2019). Copyright lessons on machine learning: What impact on algorithmic art? *10 JIPITEC*, 398–411 [in English]
- Clark, C. (1996). The Future of Copyright in a Digital Environment. *Kluwer Law International*, 139–146[in English]
- De la Durantaye, K. (2023). Garbage in, garbage out. Regulating generative AI through copyright law. *ZUM 10/2023*, 645–660. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4572952](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4572952) [in English]
- Directive 2019/790/EU on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC of 17 April 2019. [https://zakon.rada.gov.ua/laws/show/984\\_022-19#Text](https://zakon.rada.gov.ua/laws/show/984_022-19#Text) (accessed: January 20, 2025) [in Ukraine]
- Dornis, T. W., & Stober, S. (2024). Urheberrecht und Training generativer KI–Modelle: Technologische und juristische Grundlagen [Copyright and training of generative AI models: Technological and legal fundamentals] *Series: Recht und Digitalisierung | Digitization and the Law – NOMOS Verlag* (Vol. 19), 1st edition, 217 p. [in German]
- Dregelies, M. (2024). KI-Training unter dem AI Act [AI training under the AI Act] *GRUR*, 1484–1495 [in German]
- European Parliament and of the Council (2024, June 13) Regulation (EU) 2024/1689 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) (Text with EEA relevance)



- <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689> (accessed: January 21, 2025) [in English]
- Graves, A. (2013). Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850* [in English]
- Griffiths, J., Synodinou, T., & Xalabarder, R. (2023). Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of Directive (EU) 2019/790 on Copyright in the Digital Single Market. *72 GRUR Int.*, 22 [in English]
- Hamann, H. (2024). Artificial intelligence and the law of machine-readability: A review of human-to-machine communication protocols and their (in)compatibility with Article 4(3) of the Copyright DSM Directive. *15 (2024) JIPITEC*, 102 [in English]
- Hofmann, F. (2016). Grundsatz der Technikneutralität im Urheberrecht? Zugleich Gedanken zu einem more technological approach [Principle of technology neutrality in copyright law? Thoughts on a more technological approach] *Jahrgang 8 (2016) / Heft 4*, 482–512 [in German]
- Hofmann, F. (2024). Zehn Thesen zu Künstlicher Intelligenz (AI) und Urheberrecht [Ten theses on artificial intelligence (AI) and copyright law] *Zeitschrift für Wettbewerb in Recht und Praxis*, 1, 11–18 [in German]
- Jagielski, M., Thakkar, O., Tramer, F., Ippolito, D., Lee, K., Carlini, N., Wallace, E., Song, S., Thakurta, A., Papernot, N., Zhang, C. (2023). Measuring forgetting of memorized training examples. *arXiv:2207.00099v2 [cs.LG]*, 9 May [in English]
- Kraetzig, V. (2024). KI-Kunst als schöpferische Zerstörung ([AI art as creative destruction]). *Neue Juristische Wochenschrift (NJW)*, 697–702 [in German]
- Masing, J. (2020). Maintaining balance. *Journal for Legal Policy (ZRP)*, 194–198
- Proposal for a Regulation of the European Parliament and the Council establishing harmonized rules on artificial intelligence (EU AI Act) amending certain legal acts of the Union. [https://eur-lex.europa.eu/resource.html?uri=cellare0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC\\_1&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellare0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF) (accessed: January 28, 2025) [in Ukraine]
- Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. Available at: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> [in English]
- Schricker, G. (2006). *Urheberrecht: Kommentar* [Copyright: Commentary] C. H. BECK Verlag (3rd ed.), 2658 p [in German]
- Senftleben, M. (2023). Generative AI and author remuneration. *54 IIC*, 1535 [in English]
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and efficient foundation language models. Retrieved from: <https://arxiv.org/pdf/2302.13971> [in English]
- Vaswani, A., Shazeer, N., & others. (2023). Attention is all you need. *arXiv:1706.03762v7 [cs.CL]*, 1–15 [in English]

- Vesala, J. (2023). Developing artificial intelligence-based content creation: Are EU copyright and antitrust law fit for purpose? *54 IIC*, 351 [in English]
- Zhornokui, Y. (2024) Comparative analysis of legal Personality of a legal Entity and artificial Intelligence within the Impact on Corporate Governance. *C. P. Journal*, 3. <https://doi.org/10.69724/2786-8834-2024-3-3-98-112> [in Ukrainian].

*Sergiy Glotov*  
*PhD in Law*  
*Researcher at the IP Scientific Research Institute*  
*of the National Academy of Law Sciences of Ukraine*  
*Charkiv, Ukraine*  
*Attorney-at-Law*

**Glotov Sergiy. Copyright v. Artificial Intelligence: Machine Learning is Not a Copyright use of a Work**

**Abstract**

The article is devoted to the issue of machine learning of artificial intelligence in terms of assessing such a primarily technical process from the point of view of copyright, namely, whether it should be considered that in this case the use of works is carried out. Given that such machine learning is usually carried out at the expense of protected works, as well as given the ambiguity of the attitude to this issue in society, which, on the one hand, is manifested in high-profile lawsuits against AI developers such as *The New York Times v. OpenAI & Microsoft Inc.* and, on the other hand, in various social initiatives aimed at promoting the development of AI tools and deepening public understanding of the benefits of AI, the issue under consideration is extremely important and interesting. The article focuses on the latest trends in the field of legislative initiatives with an emphasis on the law of the European Union. In particular, a brief description is given of the recently adopted Regulation (EU) 2024/1689 of June 13, 2024, which is also known as the EU Artificial Intelligence Law. It is noted that in the case of AI training using works that have passed into the public domain, the output results of AI will correspond to the level of knowledge of a century ago, i.e., will be irrelevant. In addition, old works can only be used in machine learning if they are digitized, which can only be done manually. This will slow down the development of the latest technologies in every sense. It is substantiated that what happens during AI machine learning is nothing more than «reading»: a special form of use of a work to which the author's right does not apply. It is established that the use of a work for the purpose of machine learning is subject to the restriction on in-depth analysis of text and data regulated by Article 4 of Directive 2019/790/EU. It is proved that cultural creativity is unthinkable without the use of previous achievements of other authors. Therefore, since the output of AI is not a copyrightable object, in the context of AI machine learning, there are no actions significant from the point of view of copyright.

**Keywords:** *artificial intelligence, copyright, machine learning, AI output, technology.*