

УДК 34**Использование объектов интеллектуальной собственности для обучения искусственного интеллекта: юридические ограничения и возможности****Жорнист Яков Михайлович**

Студент,
Национальный исследовательский
университет «Высшая школа экономики»,
101000, Российская Федерация, Москва, ул. Мясницкая, 20;
e-mail: yashazhor@gmail.com

Аннотация

Целью данного исследования является анализ правовых аспектов использования объектов интеллектуальной собственности в процессе обучения искусственного интеллекта, а также выявление возможных конфликтов между развитием технологий и соблюдением авторских прав. Методология включает сравнительно-правовой анализ подходов к регулированию машинного обучения и майнинга данных (TDM) в различных юрисдикциях: Европейском союзе, США и Японии. В статье раскрываются две основные модели регулирования: модель добросовестного использования (fair use), характерная для США и Японии, и модель opt-out/opt-in, реализованная в ЕС. По результатам исследования автор делает вывод о большей эффективности модели opt-out, обеспечивающей баланс между интересами правообладателей и пользователей. Также в работе подчеркнута необходимость развития специализированных цифровых платформ, способствующих формированию легального рынка данных. В заключении отмечается важность выработки международных стандартов регулирования, способных одновременно защищать интеллектуальную собственность и стимулировать развитие инноваций в области искусственного интеллекта.

Для цитирования в научных исследованиях

Жорнист Я.М. Использование объектов интеллектуальной собственности для обучения искусственного интеллекта: юридические ограничения и возможности // Вопросы российского и международного права. 2025. Том 15. № 4А. С. 158-167.

Ключевые слова

Интеллектуальная собственность, авторское право, искусственный интеллект, машинное обучение, майнинг данных, регулирование, коммерческое использование.

Введение

Искусственный интеллект, как область науки, техники и высоких технологий, на протяжении последних лет неизменно находится в центре пристального внимания как академического сообщества, так и государственных структур на международном уровне [European Parliament and Council, 2024]. Особое внимание уделяется вопросам, связанным с возможными нарушениями авторских прав в ходе разработки и применения ИИ. Стремительное развитие искусственного интеллекта, сопровождающееся быстрым внедрением новых технологий и алгоритмов, приводит к возникновению ряда фундаментальных юридических и этических вопросов, прежде всего связанных с соблюдением законодательства об интеллектуальной собственности, включая, в первую очередь, вопросы соблюдения авторских прав на входные данные, используемые для обучения систем ИИ. Системные исследования и консультации по этой теме ведут, в частности, Бюро по авторскому праву США в рамках инициативы “Copyright and Artificial Intelligence” и Британское ведомство интеллектуальной собственности [UK Intellectual Property Office, 2024; U.S. Copyright Office, 2025].

Первым и наиболее важным вопросом, который поднимается в настоящей статье, является вопрос о том, нарушает ли искусственный интеллект авторские права на объекты, если он использует и обрабатывает такие данные исключительно в целях обучения, то есть для совершенствования собственных алгоритмов и повышения точности прогнозирования. Это становится особенно актуальным в контексте растущего числа ИИ-приложений, работающих с текстами, изображениями и другими видами интеллектуальной собственности. Вторым, также не менее важным аспектом, требующим комплексного рассмотрения, выступает вопрос о возможности и правомерности использования пользователем, а также создателем программного обеспечения, данных, представляющих собой объекты интеллектуальной собственности, которые были получены от правообладателя, в коммерческих интересах и с целью извлечения прибыли.

Актуальность данных вопросов обусловлена тем, что в современных условиях искусственный интеллект, включая разнообразные системы машинного обучения, уже вовлечён в создание широкого спектра услуг, продуктов и изобретений. К таким изобретениям относятся, в частности, произведения музыки, визуального и изобразительного искусства, текстовые работы, а также различные новые изображения, сгенерированные на основе алгоритмов глубокого обучения. В каждом из этих случаев возникает острая проблема, связанная с необходимостью использования больших объемов данных для полноценного и эффективного машинного обучения. Однако зачастую информация, используемая для обработки и обучения, принадлежит третьим лицам, обладающим исключительными правами на такие данные, что в итоге приводит к возникновению конфликтов между технологическим прогрессом и защитой прав интеллектуальной собственности.

Важно подчеркнуть, что законодательство в большинстве стран мира не всегда поспевает за теми быстрыми и иногда даже революционными изменениями, которые происходят в области ИИ и сопутствующих ему технологий. Такое отставание нормативного регулирования может создавать правовую неопределённость, как для разработчиков, так и для правообладателей. В результате наблюдается разнообразие подходов и видений относительно способов регулирования указанных вопросов. В настоящей статье делается попытка выработать и предложить универсальный подход, который потенциально мог бы быть применим в различных юрисдикциях и обеспечивать баланс между инновациями и необходимостью защиты прав

авторов.

Прежде всего, целесообразно начать с рассмотрения проблемы использования информации, находящейся под охраной авторского права, в целях обучения искусственного интеллекта. Для комплексного понимания вопроса, однако, необходимо кратко охарактеризовать сам процесс машинного обучения и его отличия от смежных понятий. В современных научных трудах существует множество различных определений искусственного интеллекта, но в упрощённом виде под ИИ понимается область науки и техники, занимающаяся созданием таких систем, которые способны выполнять задачи, требующие человеческого интеллекта.

Если же мы говорим об обучении искусственного интеллекта, то следует выделить специфический раздел — машинное обучение (МО), основной задачей которого выступает обучение машины мыслить, делать выводы и решать задачи на основе полученных данных. Ключевая идея здесь заключается в том, что система анализирует огромные массивы информации, извлекает из них закономерности, учится на примерах и на этой основе вырабатывает собственные выводы и предсказания, которые затем используются для дальнейшего совершенствования модели.

Вместе с тем, важно не путать машинное обучение в составе искусственного интеллекта и так называемый майнинг текстов и данных (TDM — text and data mining). Майнинг текстов и данных представляет собой отдельный процесс, суть которого заключается в анализе больших объёмов информации с использованием различных аналитических методов и специализированных алгоритмов для извлечения новых знаний и выявления скрытых закономерностей [Simplilearn, 2024]. Можно сказать, что именно данные, полученные в результате применения методов data mining, в определённой степени способствуют динамическому обучению машин, позволяя системам ИИ совершенствовать свои алгоритмы на основе новых вводных. По сути, майнинг данных выступает в качестве своеобразной предварительной стадии в процессе машинного обучения. В настоящем исследовании аббревиатура TDM будет использоваться исключительно в контексте машинного обучения и только для описания соответствующих процессов.

Правовые аспекты регулирования применения ИИ

Современные модели искусственного интеллекта, для достижения высокой степени эффективности, требуют использования значительных по объёму массивов данных, которые зачастую представляют собой целые датасеты, сформированные из объектов, защищённых авторским правом. К таким объектам относятся, в том числе, разнообразные изображения, музыкальные и литературные произведения, а также иные результаты интеллектуальной деятельности человека. Использование подобных данных, особенно без получения явного, однозначного и предварительного согласия правообладателя, неизбежно чревато возникновением серьёзных юридических и этических проблем, в первую очередь связанных с нарушением прав интеллектуальной собственности, а также возможными судебными исками со стороны правообладателей.

На сегодняшний день в большинстве стран мира отсутствует прямое и специально ориентированное законодательное регулирование, посвящённое исключительно вопросам использования объектов интеллектуальной собственности на стадии машинного обучения. В подавляющем большинстве юрисдикций применяется совокупность национальных и

международных правовых актов, регулирующих общие вопросы добычи данных (data mining), а также вопросы, связанные с защитой авторских прав. Согласно положениям большинства действующих нормативных актов, нарушение прав интеллектуальной собственности может наступить уже на самом начальном этапе — этапе сбора, обработки и последующего анализа информации для целей обучения.

Одним из наиболее известных и часто упоминаемых в данной области нормативных документов является Директива (ЕС) 2019/790 Европейского парламента и Совета от 17 апреля 2019 года “Об авторском праве и смежных правах в цифровом едином рынке”. Этот правовой акт закрепляет основные положения, регулирующие майнинг текстов и данных (TDM, text and data mining), а также вводит специальное определение этого понятия в статье 2 [Directive (EU) 2019/790 on Copyright and Related Rights in the Digital Single Market, 2019]. Согласно данной статье, под TDM понимается “любая автоматизированная аналитическая техника, направленная на анализ текстов и данных в цифровой форме с целью получения информации, которая включает, но не ограничивается выявлением закономерностей, тенденций и корреляций”. Данная формулировка подчеркивает, что спектр возможных целей и результатов анализа весьма широк и не ограничивается исключительно научными или исследовательскими задачами.

Директива устанавливает два принципиально различных подхода к регулированию использования объектов интеллектуальной собственности для целей майнинга текстов и данных, при этом выбор подхода зависит, прежде всего, от целей, в рамках которых осуществляется обработка и анализ информации. Первый подход, зафиксированный в статье 3 Директивы, предусматривает ряд исключений для исследовательских организаций, университетов, научных институтов, а также для культурных учреждений. Указанное регулирование распространяется только на те данные, которые были получены на законных основаниях, то есть легально, например, по лицензионному соглашению, либо размещены в открытом доступе в сети Интернет. Это важное условие, которое позволяет отделять законные действия от противоправных.

Второй подход, детализированный в статье 4 той же директивы, предусматривает более общий режим, согласно которому решение о разрешении или запрете использования объектов интеллектуальной собственности полностью остаётся на усмотрение правообладателя. Правообладатель может прямо и однозначно установить запрет на использование своих данных, либо, напротив, не предпринимать никаких действий, в результате чего его данные могут быть использованы в порядке, установленном законом. Если правообладатель желает ограничить возможности использования своих данных, он обязан выразить такой запрет явно, например, путём прямого указания соответствующих положений в договоре или соглашении. При этом стоит подчеркнуть, что такой режим допускает использование данных как для научных, так и для коммерческих целей, если отсутствует прямой запрет со стороны правообладателя.

Таким образом, можно прийти к выводу, что в случае, если запрет установлен, действуют только исключения, перечисленные в статье 3 директивы. Если же запрет отсутствует и никак не выражен, тогда данные могут использоваться не только для исследовательских целей, но и для иных целей, включая коммерческое использование, что реализуется посредством модели «отказа от участия» (opt-out). Такой подход позволяет пользователям действовать достаточно свободно, если только правообладатель не воспользовался своим правом на запрет.

Вместе с тем, важно отметить, что уже спустя относительно короткий промежуток времени после принятия указанной директивы в Европейском союзе начали формироваться предпосылки для создания и развития полноценного рынка по формированию и продаже

специализированных датасетов, предназначенных для целей майнинга данных [Kluwer Copyright Blog. The New Copyright Directive: Text and Data Mining (Articles 3 and 4), 2024]. Это обстоятельство объективно ведёт к тому, что правообладатель получает эффективные инструменты для полного контроля над распоряжением, использованием, а также коммерциализацией своих объектов авторского права.

Однако остаётся неурегулированным вопрос о возможности получения денежного вознаграждения за использование подобных данных, поскольку указанный вопрос прямо не затрагивается в тексте нормативного акта. Например, пункт 17 преамбулы директивы содержит запрет исключительно в отношении научных исследований, в то время как о других способах использования данных ничего не сказано. В результате эксперты делают вывод о том, что в остальных случаях вопрос подлежит регулированию на национальном уровне в соответствии с внутренним законодательством конкретного государства-члена [UK Intellectual Property Office. Copyright and Artificial Intelligence – Consultation Outcome, 2024].

Далее необходимо рассмотреть и иные подходы к регулированию использования объектов интеллектуальной собственности в других юрисдикциях, что позволит провести сравнительный анализ различных моделей и выявить их особенности. Примером может служить Япония, где ещё к 2009 году в законодательстве появились нормы, дающие определение анализу данных и текстов, а также регулирующие их использование. В соответствии с японским законом об авторском праве, в частности статьёй 30–4 [Copyright Research and Information Center. Copyright Law of Japan, 2024], разрешается использовать данные любым необходимым способом при условии соблюдения ряда критериев, среди которых — наличие конкретной, практически значимой цели, а также запрет на использование данных исключительно в развлекательных целях. Помимо этого, законодатель гарантирует правообладателю защиту от необоснованного ущерба, который может возникнуть в результате соответствующего использования его интеллектуальной собственности. Основной целью использования данных должно быть их аналитическое изучение. Закон прямо разрешает проведение анализа данных и текстов для коммерческих целей, при этом любые ограничения на такую деятельность признаются недействительными и не имеющими юридического значения. Также важно, что в японском законодательстве отсутствуют какие-либо прямые запреты или ограничения на передачу созданных копий произведений третьим лицам. Всё это позволяет заключить, что Япония занимает одну из наиболее либеральных позиций в сфере регулирования майнинга текстов и данных, если не самую либеральную среди развитых стран [Privacy World. Japan's New Draft Guidelines on AI and Copyright: Is it Really OK to Train AI Using Pirated Materials?, 2024].

В Соединённых Штатах Америки, напротив, отсутствует специально посвящённое регулирование, касающееся вопросов майнинга текстов и данных. Оценка законности соответствующих действий осуществляется в рамках рассмотрения судебных дел и базируется на доктрине добросовестного использования (*fair use*). Например, в известных делах *Authors Guild v. Google* [Authors Guild v. Google, Inc., No. 13-4829 (2d Cir. 2015), 2015] и *Authors Guild v. HathiTrust* [Authors Guild, Inc. v. HathiTrust, No. 12-4547 (2d Cir. 2014), 2014] американские суды пришли к выводу, что, если использование данных носит трансформирующий характер и не подрывает коммерческую ценность оригинального произведения на рынке, оно может быть признано добросовестным использованием. Основное назначение такого использования должно быть связано с научными исследованиями, образовательными инициативами или развитием новых технологий, что подчёркивает социальную значимость подобных процессов.

Примеры платформ, предоставляющих доступ к объектам интеллектуальной собственности для целей анализа данных и машинного обучения

В контексте регулирования доступа к объектам интеллектуальной собственности для целей анализа данных и последующего машинного обучения необходимо рассмотреть конкретные примеры современных сервисов и платформ, функционирующих в различных отраслях науки, искусства и бизнеса. Одним из наиболее ярких и широко используемых в академическом сообществе примеров является сервис “arXiv” [arXiv. arXiv Submission Licenses, 2024], представляющий собой обширный онлайн-архив научных публикаций, статей, препринтов и иных материалов. Благодаря данному сервису исследователи по всему миру получают уникальную возможность быстро и беспрепятственно находить, просматривать и скачивать необходимые для анализа тексты и публикации. Примечательно, что подавляющее большинство материалов arXiv размещается в открытом доступе и, как правило, распространяется по условиям свободных лицензий. Это означает, что на такие публикации практически не распространяются ограничения, характерные для охраняемых объектов авторского права, и исследователи могут свободно использовать, скачивать и анализировать контент для некоммерческих научно-исследовательских целей. Что касается хранения данных, то каких-либо специальных или дополнительных требований к условиям хранения не предъявляется, однако при этом, важно отметить, что настоятельно рекомендуется уважать права авторов и соблюдать положения соответствующих лицензий. Для использования материалов arXiv в коммерческих целях, как правило, требуется получение специального разрешения от правообладателя.

В такой индустрии, как кино и смежных с ней областях примечательным примером служит платформа IMDb [IMDb Developer. Non-Commercial Datasets, 2024], представляющая собой одну из крупнейших в мире баз данных, содержащую разноплановую и структурированную информацию об актёрах, телесериалах, съёмочных командах, фильмах и других объектах, относящихся к кинематографу. Платформа IMDb предоставляет своим пользователям доступ к огромному массиву информации, что открывает широчайшие перспективы для применения таких данных в аналитических, исследовательских, а также образовательных и коммерческих проектах, связанных с развитием машинного обучения и искусственного интеллекта. Следует, однако, подчеркнуть, что данные IMDb, как правило, разрешается использовать только в личных, образовательных и некоммерческих целях. Для их применения в коммерческих целях необходимо получить специальное разрешение либо заключить лицензионное соглашение с администрацией платформы или непосредственным правообладателем соответствующих данных.

Ещё одним примечательным примером платформы, заслуживающей упоминания, выступает сервис Google Books, обеспечивающий доступ к огромному количеству оцифрованных книг, их отдельных фрагментов, метаданных, а в некоторых случаях и к полному тексту произведений. Сервис Google Books предоставляет уникальную возможность поиска и изучения книг по различным параметрам, что значительно облегчает проведение анализа и позволяет использовать такие данные в рамках различных исследовательских, образовательных и аналитических проектов. Вместе с тем, многие книги, доступные на платформе, по-прежнему охраняются авторским правом, вследствие чего доступ к ним ограничивается лишь отдельными отрывками или фрагментами. Для полноценного и особенно для коммерческого использования

этих материалов требуется получение согласия правообладателя.

Проанализировав вышеизложенное, можно прийти к промежуточному выводу о том, что на современном этапе развития международного права существует две основные, наиболее востребованные и применяемые модели регулирования вопросов, связанных с использованием объектов интеллектуальной собственности для обучения искусственного интеллекта и развития технологий машинного обучения. Первая модель, типичная для таких государств, как Япония и США, основывается на доктрине так называемого добросовестного использования (*fair use*), предоставляющей определённую свободу действий при условии, что интересы правообладателя не нарушаются и отсутствует прямой, однозначно выраженный запрет. Тем не менее, данный подход сопряжён с определёнными рисками для правообладателей, поскольку оставляет широкое поле для разночтений и судебных споров, и не всегда может считаться оптимальным в плане правовой определённости.

Вторая модель, обладающая большей гибкостью и универсальностью, реализуется в Европейском союзе, где использование информации для обучения машин возможно при наличии чётко определённых законом оснований. При некоммерческом характере использования пользователю предоставляется широкая свобода действий в отношении объектов, легально введённых в гражданский оборот. Для иных видов использования действуют две модели: *opt-out* (если правообладатель не запретил явно — использование допускается) и *opt-in* (использование возможно только при получении прямого, однозначного согласия правообладателя). С точки зрения комфорта для бизнеса и удобства пользователей, *opt-out*-модель выглядит предпочтительнее и современной.

С учётом вышесказанного, представляется важным и целесообразным поддерживать и развивать современные платформы, облегчающие законное приобретение и использование данных, необходимых для обучения ИИ. Так, сегодня активно развиваются сервисы типа Amazon Web Services (AWS) Data Exchange, предназначенные для поиска, приобретения и использования сторонних данных в облачной среде AWS, где доступны как бесплатные, так и платные датасеты [Amazon Web Services. AWS Data Exchange – User Guide, 2025]. Подобные платформы в значительной степени способствуют формированию легального рынка данных и позволяют максимально учитывать как интересы правообладателей, так и потребности разработчиков и исследователей.

Заключение

В заключение отметим, что использование объектов интеллектуальной собственности в процессе обучения искусственного интеллекта и развития новых технологий машинного обучения неизбежно обостряет проблему поиска баланса между инновационным развитием и защитой авторских прав. На сегодняшний день в разных странах мира предлагаются различные модели правового регулирования: США и Япония склоняются к гибкому подходу посредством института добросовестного использования, тогда как Европейский союз использует более формализованную систему *opt-out* и *opt-in*. Пример AWS Data Exchange иллюстрирует возможности для дальнейшего развития легального доступа к данным и расширения сотрудничества между всеми участниками рынка. Постепенно формируется тенденция к выработке универсальных международных стандартов [Stieper, 2020], способных одновременно учитывать права авторов и стимулировать инновации в сфере искусственного интеллекта и смежных технологий.

Библиография

1. Amazon Web Services. *AWS Data Exchange – User Guide* [Электронный ресурс]. 2025. – URL: <https://docs.aws.amazon.com/data-exchange/> (дата обращения: 15.07.2025).
2. arXiv. *arXiv Submission Licenses* [Электронный ресурс]. 2024. – URL: <https://arxiv.org/licenses/> (дата обращения: 15.07.2024).
3. Copyright Research and Information Center. *Copyright Law of Japan* [Электронный ресурс]. 2024. – URL: <https://www.cric.or.jp/english/clj/> (дата обращения: 15.07.2024).
4. European Parliament and Council. *Directive (EU) 2019/790 on Copyright and Related Rights in the Digital Single Market* [Электронный ресурс]. 2019. – URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0790> (дата обращения: 15.07.2024).
5. European Parliament and Council. *Regulation (EU) 2024/1689 on Artificial Intelligence (AI Act)* [Электронный ресурс]. 2024. – URL: <https://eur-lex.europa.eu/eli/reg/2024/1689> (дата обращения: 15.07.2025).
6. IMDb Developer. *Non-Commercial Datasets* [Электронный ресурс]. 2024. – URL: <https://developer.imdb.com/non-commercial-datasets/> (дата обращения: 15.07.2024).
7. Justia Law. *Case: Authors Guild v. Google, Inc., No. 13-4829 (2d Cir. 2015)* [Электронный ресурс]. 2015. – URL: <https://law.justia.com/cases/federal/appellate-courts/ca2/13-4829/13-4829-2015-10-16.html> (дата обращения: 15.07.2024).
8. Justia Law. *Case: Authors Guild, Inc. v. HathiTrust, No. 12-4547 (2d Cir. 2014)* [Электронный ресурс]. 2014. – URL: <https://law.justia.com/cases/federal/appellate-courts/ca2/12-4547/12-4547-2014-06-10.html> (дата обращения: 15.07.2024).
9. Kluwer Copyright Blog. *The New Copyright Directive: Text and Data Mining (Articles 3 and 4)* [Электронный ресурс]. 2024. – URL: <https://copyrightblog.kluweriplaw.com/> (дата обращения: 15.07.2024).
10. Privacy World. *Japan's New Draft Guidelines on AI and Copyright: Is it Really OK to Train AI Using Pirated Materials?* [Электронный ресурс]. 18.03.2024. – URL: <https://www.privacyworld.blog/2024/03/japans-new-draft-guidelines-on-ai-and-copyright-is-it-really-ok-to-train-ai-using-pirated-materials/> (дата обращения: 15.07.2025).
11. Simplilearn. *Data Mining vs Machine Learning: Key Differences* [Электронный ресурс]. 2024. – URL: <https://www.simplilearn.com/data-mining-vs-machine-learning-article> (дата обращения: 15.07.2024).
12. Stieper M. The Relationship of the Mandatory Exceptions of the DSM Directive to the Optional Exceptions of the InfoSoc Directive // *GRUR*. 2020. № 1.
13. UK Intellectual Property Office. *Copyright and Artificial Intelligence – Consultation Outcome* [Электронный ресурс]. 2024. – URL: <https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence/> (дата обращения: 15.07.2025).
14. U.S. Copyright Office. *Copyright and Artificial Intelligence* [Электронный ресурс]. 2025. – URL: <https://www.copyright.gov/ai/> (дата обращения: 15.07.2025).
15. World Intellectual Property Organization. *IP Policy Toolkit: Getting the Innovation Ecosystem Ready for AI* [Электронный ресурс]. 2024. – URL: <https://www.wipo.int/publications/en/details.jsp?id=4572> (дата обращения: 15.07.2025).

Use of Intellectual Property Objects for Artificial Intelligence Training: Legal Constraints and Opportunities

Yakov M. Zhornist

Student,
National Research University "Higher School of Economics",
101000, 20, Myasnitskaya str., Moscow, Russian Federation;
e-mail: yashazhor@gmail.com

Abstract

This study aims to analyze the legal aspects of using intellectual property objects for artificial intelligence training and identify potential conflicts between technological development and copyright compliance. The methodology involves a comparative legal analysis of approaches to

regulating machine learning and text/data mining (TDM) across different jurisdictions: the European Union, United States, and Japan. The article examines two primary regulatory models: the fair use doctrine characteristic of the US and Japan, and the opt-out/opt-in model implemented in the EU. The research concludes that the opt-out model proves more effective in balancing rights holders' and users' interests. The study also emphasizes the need to develop specialized digital platforms that facilitate the creation of legal data markets. The findings highlight the importance of establishing international regulatory standards capable of simultaneously protecting intellectual property and fostering innovation in artificial intelligence development.

For citation

Zhornist Ya.M. (2025) Ispolzovanie obektov intellektualnoy sobstvennosti dlya obucheniya iskusstvennogo intellekta: yuridicheskie ogranicheniya i vozmozhnosti [Use of Intellectual Property Objects for Artificial Intelligence Training: Legal Constraints and Opportunities]. *Voprosy rossiiskogo i mezhdunarodnogo prava* [Matters of Russian and International Law], 15 (4A), pp. 158-167.

Keywords

Intellectual property, copyright, artificial intelligence, machine learning, data mining, regulation, commercial use.

References

1. Amazon Web Services. AWS Data Exchange – User's Guide [Electronic resource]. 2025. – URL: <https://docs.aws.amazon.com/data-exchange/> / (date of access: 07/15/2025).
2. arXiv. Licenses for sending the arXiv [Electronic resource]. 2024. – URL: <https://arxiv.org/licenses/> / (date of access: 07/15/2024).
3. Center for Research and Information in the field of Copyright. The Law of Japan on copyright [Electronic resource]. 2024. – URL: <https://www.cric.or.jp/english/clj/> / (date of access: 07/15/2024).
4. The European Parliament and the Council. Directive (EU) 2019/790 on copyright and related rights in the single digital market [Electronic resource]. 2019. – URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0790> (accessed: 07/15/2024).
5. The European Parliament and the Council. Regulation (EU) 2024/1689 on Artificial Intelligence (AI Act) [Electronic resource]. 2024. – URL: <https://eur-lex.europa.eu/eli/reg/2024/1689> (date of request: 07/15/2025).
6. IMDb developer. Non-commercial datasets [Electronic resource]. 2024. – URL: <https://developer.imdb.com/non-commercial-datasets/> / (date of application: 07/15/2024).
7. The Law on Justice. Case: Guild of Authors v. Google, Inc., No. 13-4829 (2nd quarter 2015) [Electronic resource]. 2015. – URL: <https://law.justia.com/cases/federal/appellate-courts/ca2/13-4829/13-4829-2015-10-16.html> (date of request: 07/15/2024).
8. The Law on Justice. Case: Guild of Authors, Inc. v. Khathitrust, No. 12-4547 (2nd quarter 2014) [Electronic resource]. 2014. – URL: <https://law.justia.com/cases/federal/appellate-courts/ca2/12-4547/12-4547-2014-06-10.html> (accessed: 07/15/2024).
9. Kluwer author's blog. The new directive on copyright: intellectual analysis of text and data (Articles 3 and 4) [Electronic resource]. 2024. – URL: <https://copyrightblog.kluweriplaw.com/> / (date of access: 07/15/2024).
10. The world of privacy. Draft of Japan's new AI and Copyright Guidelines: Is it really possible to train AI using pirated materials? [electronic resource]. 03/18/2024. – URL: <https://www.privacyworld.blog/2024/03/japans-new-draft-guidelines-on-ai-and-copyright-is-it-really-ok-to-train-ai-using-pirated-materials/> / (date of access: 07/15/2025).
11. Simple training. Data mining versus machine learning: key differences [Electronic resource]. 2024. – URL: <https://www.simplilearn.com/data-mining-vs-machine-learning-article> (date of request: 07/15/2024).
12. Stiper M. The relationship of mandatory exceptions from the DSM Directive with optional exceptions from the InfoSoc Directive // GRUR. 2020. No. 1.
13. UK Intellectual Property Office. Copyright and artificial intelligence – the results of consultations [Electronic resource]. 2024. – URL: <https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence/> / (date of access: 07/15/2025).

-
14. US Copyright Office. Copyright and artificial intelligence [Electronic resource]. 2025. – URL: <https://www.copyright.gov/ai/> (date of access: 07/15/2025).
 15. World Intellectual Property Organization. IP policy toolkit: Preparing an innovative ecosystem for the introduction of AI [Electronic resource]. 2024. – URL: <https://www.wipo.int/publications/en/details.jsp?id=4572> (date of request: 07/15/2025).