

ПРАВО

УДК 347.78

ИСПОЛЬЗОВАНИЕ ПРОИЗВЕДЕНИЙ, ОХРАНЯЕМЫХ АВТОРСКИМ ПРАВОМ, ДЛЯ ОБУЧЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: НОВЫЙ ЭТАП РАЗВИТИЯ ДОКТРИНЫ ДОБРОСОВЕСТНОГО ИСПОЛЬЗОВАНИЯ

А. Н. Артемова

Институт философии и права СО РАН (г. Новосибирск)
artemova-an-1991@yandex.ru

Аннотация. Статья посвящена анализу знакового определения Окружного суда США (Северный округ Калифорнии) по делу Bartz et al. v. Anthropic PBC (2025), ставшего первым прецедентом всесторонней оценки применения доктрины добросовестного использования (fair use) в отношении практики обучения больших языковых моделей на материалах, охраняемых авторским правом. На основе детального разбора аргументации суда автор демонстрирует, где проходит граница правомерности использования данных: использование объектов авторского права для обучения искусственного интеллекта и создание цифровой библиотеки путем оцифровки легально приобретенных книг было признано добросовестным использованием, в то время как использование пиратских копий для создания такой библиотеки – недобросовестным использованием. Целью данной статьи является анализ ключевых правовых позиций, сформулированных судом в данном определении, их значения для понимания эволюции доктрины добросовестного использования в цифровую эпоху и практических последствий для разработчиков искусственного интеллекта и авторов произведений.

Ключевые слова: доктрина добросовестного использования, авторское право, искусственный интеллект, большие языковые модели, машинное обучение, преобразующий характер.

Для цитирования: Артемова, А. Н. (2025). Использование произведений, охраняемых авторским правом, для обучения искусственного интеллекта: новый этап развития доктрины добросовестного использования. *Respublica Literaria*. Т. 6. № 3. С. 136-143. DOI: 10.47850/RL.2025.6.3.136-143.

USING COPYRIGHTED WORKS TO TRAIN ARTIFICIAL INTELLIGENCE: A NEW STAGE IN THE DEVELOPMENT OF THE FAIR USE DOCTRINE

A. N. Artemova

Institute of Philosophy and Law SB RAS (Novosibirsk)
artemova-an-1991@yandex.ru

Abstract. The article analyzes the landmark order of the US District Court (Northern District of California) in the case of Bartz et al. v. Anthropic PBC (2025), which became the first precedent for a comprehensive assessment of the application of the fair use doctrine in relation to the practice of training large language models on copyrighted materials. Based on a detailed analysis of the court's reasoning, the author demonstrates where the line of legality for using data lies: using copyrighted objects to train artificial intelligence and to create a digital library by digitizing legally acquired books was recognized as fair use, while using pirated copies to create such a library was considered unfair use. The purpose of this article is to analyze the key legal positions formulated by the court in this order, their significance for understanding the evolution of the fair use doctrine in the digital age and the practical consequences for artificial intelligence developers and authors of works.

Keywords: fair use doctrine, copyright, artificial intelligence, large language models, machine learning, transformative nature.

For citation: Artemova, A. N. (2025). Using Copyrighted Works to Train Artificial Intelligence: A New Stage in the Development of the Fair use Doctrine. *Respublica Literaria*. Vol. 6. No. 3. Pp. 136-143. DOI: 10.47850/RL.2025.6.3.136-143.

Стремительное развитие генеративного искусственного интеллекта (далее – ИИ) требует обучения больших языковых моделей (LLM) на колоссальных массивах текстовых данных. Лучшими источниками таких данных признаны книги [Артемова, 2024, с. 185]. Однако авторско-правовая охрана литературных произведений накладывает определенные ограничения на их использование без согласия правообладателей. Это ставит перед правом фундаментальный вопрос: при каких условиях использование охраняемых авторским правом произведений для обучения ИИ является законным?

Для ответа на это вопрос в США используется доктрина добросовестного использования (fair use), впервые сформулированная в 1841 г. в деле *Folsom v. Marsh*¹. В упрощенном виде доктрина предусматривает возможность свободного использования объектов авторского права, если такое использование является разумным, что позволяет быстро адаптироваться к развитию технологий, появлению новых способов использования произведений, потребностям общества и т. д. [Калятин, 2013, с. 137]. Первоначально доктрина существовала в рамках «общего права» (common law), позднее получила закрепление в законе [Луткова, 2016, с. 187].

Каждый случай предполагаемого добросовестного использования анализируется судом индивидуально на основе четырех факторов, перечисленных в статье 107 Закона США об авторском праве 1976 г.²:

- 1) цель и характер использования (является ли использование коммерческим или некоммерческим; является ли использование «преобразующим»);
- 2) природа охраняемого произведения (насколько высок уровень творчества в использованном произведении);
- 3) объем и существенность использованной части (было ли использовано небольшое количество материала по отношению ко всему произведению; была ли использована существенная часть – самая важная, узнаваемая или ценная часть произведения);
- 4) влияние использования на потенциальный рынок или стоимость произведения (наносит ли использование ущерб текущему или потенциальному рынку оригинального произведения, заменяет ли покупку оригинала; может ли такое использование подорвать новые рынки для правообладателя).

Таким образом, fair use – это гибкая доктрина, позволяющая судам на основе анализа четырех факторов разрешать ограниченное, социально полезное использование чужого авторского контента без разрешения, когда это служит целям свободы слова, образования, критики, инноваций и прогресса, не нанося при этом неоправданного ущерба интересам правообладателя.

¹ *Folsom v. Marsh*, 9 F. Cas. 342. (1841). [Online]. Available at: <https://law.justia.com/cases/federal/district-courts/massachusetts/madce/9fcas342/4104271/220/no-4.html> (Accessed: 05 August 2025).

² U. S. Code Title 17 – Copyrights. Legal Information Institute. Cornell Law School. [Online]. Available at: <https://www.law.cornell.edu/uscode/text/17/> (Accessed: 05 August 2025).

Основой, «сердцем» доктрины добросовестного использования выступает идея преобразующего использования, добавляющего новую ценность к оригиналу. Преобразующим признается такое использование, которое добавляет новое значение, сообщение, функцию или выражение к исходному произведению (например, пародия, критика, комментарий, создание мема, цитирование в научной работе), обладает свойством «трансформации». «Свойство трансформации подтверждается в том случае, если будет установлено, что использование копируемого произведения соответствует цели Закона 1976 г. – стимулированию творческого обогащения общества» [Луткова, 2016, с. 187]. Чем более преобразующим является использование, тем больше вероятность признания его добросовестным. Если использование преобразует оригинал, придавая ему новое значение, это сильно склоняет чашу весов в пользу признания его добросовестным, даже если другие факторы (например, коммерческий характер) говорят против [Ворожевич, 2019, с. 45].

Как справедливо отмечает В. О. Калятин, «законодательство об авторском праве находится в начале нового этапа своего развития – этапа глобального использования произведений в цифровой форме» [Калятин, 2013, с. 137]. Как следствие, доктрина добросовестного использования должна адаптироваться к новым потребностям общества – потребностям в развитии технологии ИИ.

23 июня 2025 г. Окружной суд США (Северного округа Калифорнии) вынес определение в порядке упрощенного производства по делу Bartz et al. v. Anthropic PBC, ставшее первым судебным решением, в котором дан ответ на вопрос о том, является ли добросовестным использование произведений, охраняемых авторским правом, для обучения искусственного интеллекта.

Фабула дела Bartz et al. v. Anthropic PBC была такова. Ответчик – компания Anthropic, основанная в январе 2021 г. бывшими сотрудниками OpenAI, разрабатывает программное обеспечение на основе искусственного интеллекта. Ключевой продукт компании – ИИ-сервис Claude, анализирующий текстовые запросы пользователей и генерирующий ответы, имитирующие человеческое общение. Первая публичная версия Claude была представлена в марте 2023 г., за ней последовало шесть обновлений. Хотя базовое использование Claude бесплатно, компания получает основной доход от платных подписок.

Возможности Claude обеспечиваются крупными языковыми моделями, обученными на обширной корпоративной цифровой библиотеке текстовых данных. С целью создания такой библиотеки Anthropic в период с 2021 по 2022 гг. осуществила массовую загрузку нескольких миллионов охраняемых авторским правом книг в цифровом формате с пиратских интернет-ресурсов (Books3, LibGen, PiLiMi). Впоследствии (в 2024 г.) Anthropic стала приобретать печатные книги (включая б/у), с которых снимала переплеты, сканировала каждую страницу и сохраняла полученные цифровые копии в формате, доступном для полнотекстового поиска. После чего оригинальные печатные экземпляры уничтожались.

Именно из этой сформированной цифровой библиотеки компания впоследствии отбирала различные наборы оцифрованных книг для обучения разрабатываемых ею крупных языковых моделей, которые являются основой ИИ-сервиса Claude. Среди использованных без разрешения правообладателей произведений оказались книги истцов. Авторы, подавшие иск о нарушении авторских прав: Андреа Бартц, Чарльз Гребер и Кирк Уоллес Джонсон.

В рамках упрощенного судебного производства ключевой вопрос заключался в следующем: подпадает ли любое из способов использования спорных произведений компанией под критерии добросовестного использования (fair use) в соответствии со статьей 107 Закона об авторском праве США. Для анализа суд выделил следующие способы использования, которые рассмотрел в отдельности:

- 1) использование произведений для обучения ИИ;
- 2) создание цифровой библиотеки путем оцифровки легально приобретенных книг;
- 3) создание цифровой библиотеки на основе пиратских копий.

Суд установил, что цель и характер *использования произведений для обучения ИИ* носят в высшей степени преобразующий характер. Цель обучения – выявление статистических закономерностей и паттернов для генерации оригинального текста – носит принципиально иной характер по сравнению с выразительной целью оригинальных произведений.

Что касается характера произведения, тот факт, что ответчик использовал художественную литературу, содержащую элементы выразительности и охраняемую авторским правом в максимальной степени, свидетельствует не в пользу добросовестного использования. Однако данный фактор не является ключевым и не перевешивает трансформативную природу использования.

Анализ объема и существенности использованной части показал, что для обучения использовались произведения истцов целиком. Однако копирование полных текстов книг было признано судом необходимым для достижения заявленной трансформативной цели – эффективного обучения больших языковых моделей, требующего огромных объемов данных для выявления лингвистических паттернов. Произведения истцов не использовались как обычные книги, а результаты Claude, получаемые на выходе, не воспроизводят оригиналы.

Наконец, анализируя четвертый фактор (влияние на потенциальный рынок/ценность произведения), суд не обнаружил доказательств того, что выводы, генерируемые Claude, прямо конкурируют с оригинальными произведениями истцов или замещают их на рынке. Поскольку фильтры Anthropic предотвращают генерацию точных копий или производных работ, Claude не воспроизводит охраняемые элементы конкретных произведений. Копии произведений, использованные для обучения конкретных LLM, не вытесняли и, согласно имеющимся данным, не будут вытеснять спрос на оригинальные произведения истцов. Усвоение общих принципов (грамматика, композиция) из тысяч текстов – закономерно точно так же, как и человеческое обучение. Требовать лицензий за использование книги в процессе обучения (как человеком, так и ИИ) абсурдно: это подорвало бы саму идею образования и творчества.

Основной же аргумент истцов о том, что обучение LLM якобы спровоцирует «взрывной рост» работ, конкурирующих с их произведениями, например, путем генерации альтернативных кратких изложений реальных событий или альтернативных образцов прозы о вымышленных событиях, был отклонен судом. Суд подчеркнул, что подобная жалоба по своей сути аналогична возражению против обучения школьников грамотному письму из опасений, что это увеличит число потенциальных конкурентов-авторов. Такого рода опасения не относятся к сфере интересов, охраняемых Законом об авторском праве. Цель закона – стимулировать создание оригинальных авторских работ, а не защищать отдельных авторов от конкуренции как таковой.

Представляет интерес аргумент истцов о том, что Anthropic лишила их возможности лицензировать свои произведения для узкоспециальной цели обучения систем ИИ. Anthropic, в свою очередь, возражает, что транзакционные издержки таких лицензионных соглашений превысили бы любую потенциальную выгоду для компании, что могло бы вынудить ее либо отказаться от работы с правообладателями, либо вовсе прекратить разработку технологии. Хотя негативное воздействие на рынок лицензий, когда использование произведения дестимулирует заинтересованных лиц к получению лицензии у правообладателя, относится к факторам, препятствующим признанию такого использования добросовестным [Ворожевич, 2019, с. 47], в данном случае суд встал на сторону ответчика. Суд указал, что, хотя потенциально такой рынок действительно мог бы развиваться, Закон об авторском праве не предоставляет авторам исключительного права контролировать такого рода «мета-использование» (использование произведений в качестве «сырья» для развития технологий, а не для их традиционного потребления).

Таким образом, суд пришел к выводу, что использование охраняемых авторским правом книг для непосредственного обучения больших языковых моделей на легально приобретенных копиях подпадает под доктрину добросовестного использования.

Далее предметом исследования выступило *создание цифровой библиотеки путем оцифровки легально приобретенных книг*. В результате анализа первого фактора суд пришел к выводу о том, что цель и характер использования произведения были преобразующими. Замена формата (из печатного в цифровой) для удобства хранения и поиска внутри корпоративной библиотеки признана трансформативной. Суд опирался на доктрину первой продажи, дающую владельцу законной копии право распоряжаться ею. Так как Anthropic приобрела свои печатные экземпляры легально, с момента покупки она получила право распоряжаться каждой копией по своему усмотрению. Каждая купленная печатная копия была скопирована для экономии места и обеспечения возможности поиска в цифровом формате, а печатный оригинал был уничтожен. В деле также отсутствовали данные о том, что Anthropic распространяла цифровые копии за пределами своей компании. В данном случае имела место замена формата (физической формы выражения произведения) без изменения содержания произведения. Таким образом, изменение формата Anthropic с печатных библиотечных копий на цифровые являлось преобразовательным в соответствии с первым фактором добросовестного использования.

При этом суд отверг аргумент Anthropic о том, что оцифровка для создания цифровой библиотеки была неотъемлемой частью использования произведений для обучения LLM и, следовательно, преобразующим использованием. Вместо этого суд указал, что само по себе преобразование печатной книги в цифровой файл для экономии места и обеспечения возможности поиска было преобразующим уже по этой причине. Следовательно, к цифровой копии следует относиться так же, как если бы купленная печатная копия была размещена в библиотеке.

Второй фактор, как и в случае с использованием для обучения, не свидетельствует в пользу добросовестного использования, но не является решающим сам по себе.

В отношении третьего фактора суд признал копирование полного текста необходимым для достижения заявленной функциональной цели – создания полноценной цифровой замены физического экземпляра.

Анализ четвертого фактора показал отсутствие ущерба для истцов в результате оцифровки печатных книг, поскольку оригинальные печатные копии уничтожались после оцифровки, а цифровые копии использовались исключительно внутри компании без распространения третьим лицам или создания общедоступной базы. В результате оцифровка печатных книг с произведениями истцов не привела к созданию дополнительных общедоступных экземпляров, способных заместить оригиналы на рынке.

Таким образом, оцифровка легально приобретенных печатных книг для внутренних корпоративных целей (архивация, поиск) при соблюдении условий уничтожения оригиналов и отсутствия внешнего распространения защищена доктриной добросовестного использования как функциональное преобразование формата (format shifting).

Напротив, попытка Anthropic обосновать *создание цифровой библиотеки из пиратских копий* книг как добросовестное использование только на том основании, что некоторые из них в конечном итоге будут использованы для обучения больших языковых моделей, была отклонена судом по следующим причинам.

Исследуя первый фактор, суд установил, что создание пиратской библиотеки без оплаты является прямой коммерческой заменой легальному рынку книг и лицензий. Цель создания самой библиотеки (накопление контента) не была признана трансформативной.

Как и в двух предыдущих случаях второй фактор говорит против признания использования добросовестным: использованные произведения (книги) являются творческими и опубликованными, что усиливает их авторско-правовую охрану.

Что касается третьего фактора, практика массового копирования миллионов книг целиком, без законных оснований, была признана чрезмерной и не соответствующей принципам добросовестного использования.

Наконец, анализ влияния на рынок показал, что создание и использование такой библиотеки наносит значительный ущерб правообладателям, лишая их доходов от продаж книг и заключения лицензионных соглашений, ведь каждая украденная копия напрямую замещает легальную покупку. Если бы такая практика стала нормой, это разрушило бы издательский рынок.

Принципиально важное значение имеет позиция суда о том, что скачивание исходных копий с пиратских сайтов не может быть оправдано последующим добросовестным использованием. Fair use не оправдывает первоначальную кражу контента, даже если позже он используется в трансформативных целях (обучение ИИ). Не в пользу Anthropic сыграло также и то, что компания отказалась раскрыть информацию о составе наборов данных для обучения ИИ и продолжала хранить неиспользуемые копии произведений «на всякий случай».

По результатам оценки четырех факторов суд установил, что создание и хранение пиратской цифровой библиотеки не является добросовестным использованием независимо от того, в каких целях использовались нелегальные копии произведений из такой библиотеки³.

Определение по делу Bartz et al. v. Anthropic PBC знаменует собой новый этап в развитии доктрины добросовестного использования, представляя собой адаптацию к вызовам эпохи искусственного интеллекта, и имеет важное значение для развития отрасли

³ Bartz et al. v. Anthropic PBC. United States District Court. Northern District of California. Order on Fair Use (2025). [Online]. Available at: https://storage.courtlistener.com/recap/gov.uscourts.cand.434709/gov.uscourts.cand.434709.231.0_4.pdf (Accessed: 05 August 2025).

ИИ. Судом предложен сбалансированный подход, признающий инновационный и трансформативный характер непосредственно обучения LLM на авторских материалах, но устанавливающий четкие правовые границы на этапе сбора и первоначального хранения данных. Ключевой вывод определения суда можно сформулировать следующим образом: технологический прогресс (ИИ) не отменяет необходимость уважать права авторов произведений на этапе сбора данных.

Выводы, сформулированные в определении, можно считать своего рода инструкцией для индустрии ИИ: обучение ИИ на произведениях соответствует критериям доктрины добросовестного использования при условии законности полученного доступа к ним. Участники отрасли должны четко понимать, что использование пиратского контента недопустимо и влечет серьезную ответственность, независимо от трансформативности конечной цели. Создание цифровых библиотек для обеспечения ИИ обучающими наборами данных является добросовестным использованием при соблюдении следующих условий: законность приобретения экземпляров произведений, уничтожение печатного оригинала после оцифровки, отсутствие последующего распространения цифровых копий.

Несмотря на то, что доктрина добросовестного использования применяется только в США, в странах романо-германской правовой семьи существует ее аналог – перечневый подход к определению случаев, когда произведения могут быть свободно использованы без согласия правообладателя и без выплаты ему вознаграждения. Оба подхода схожи по своей сути: «... как доктрина добросовестного использования, так и конкретный перечень случаев свободного использования устанавливают содержательные границы исключительного права – способы использования объекта авторского права, на которые не распространяется господство правообладателя» [Ворожевич, 2019, с. 44].

В России перечень случаев свободного использования установлен в статьях 1273-1279 Гражданского кодекса РФ. Однако ни один из случаев свободного использования произведений, установленных законом, не предусматривает возможности использования произведений для обучения ИИ. При этом устранение необоснованных нормативно-правовых ограничений для разработки отечественных больших генеративных моделей (в том числе создание возможностей для обучения больших генеративных моделей на больших массивах информации) является одним из основных направлений создания комплексной системы нормативно-правового регулирования общественных отношений, связанных с развитием и использованием технологий искусственного интеллекта, и обеспечения безопасности применения таких технологий в РФ⁴.

В этой связи принципиально важным является дополнение случаев свободного использования произведений, охраняемых авторским правом, таким способом использования, как интеллектуальный анализ, в целях обучения искусственного интеллекта. Данная новелла будет соответствовать как общемировым тенденциям развития авторского права, так и особенностям национальной правовой системы; отвечать потребностям развития технологии искусственного интеллекта и национальным интересам Российской Федерации.

⁴ Указ Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации». Национальная стратегия развития искусственного интеллекта на период до 2030 года (в редакции Указа Президента Российской Федерации от 15.02.2024 № 124). П. 175. [Электронный ресурс]. Президент России. URL: <http://www.kremlin.ru/acts/bank/44731/page/2> (дата обращения: 18.07.2024).

Список литературы / References

Артемова, А. Н. (2024). Использование объектов авторского права в машинном обучении: в поисках баланса интересов правообладателей и общества. *Respublica Literaria*. Т. 5. № 3. С. 184-194.

Artemova, A. N. (2024). Use of Copyright Works in Machine Learning: Search for a Balance of Interests of Copyright Holders and Society. *Respublica Literaria*. Vol. 5. No. 3. Pp. 184-194. (In Russ.)

Ворожевич, А. С. (2019). Доктрина добросовестного использования в сфере авторского права. *Хозяйство и право*. № 10. С. 41-61.

Vorozhevich, A. S. (2019). The Doctrine of Fair Use in the Sphere of Copyright. *Business and Law*. No. 10. Pp. 41-61. (In Russ.)

Калятин, В. О. (2013). О некоторых тенденциях в развитии концепции «добросовестного использования» в современном авторском праве стран «общего права». *Право. Журнал Высшей школы экономики*. № 3. С. 136-150.

Kalyatin, V. O. (2013). On Some Trends in the Development of the Concept of Fair Use in Modern Copyright Law of Common Law Countries. *Law. Journal of the Higher School of Economics*. No. 3. Pp. 136-150. (In Russ.)

Луткова, О. В. (2016). Доктрина добросовестного использования произведений в современном авторском праве США. *Право. Журнал Высшей школы экономики*. № 2. С. 186-199.

Lutkova, O. V. (2016). The Fair use Doctrine in Modern US Copyright Law. *Law. Journal of the Higher School of Economics*. No. 2. Pp. 186-199. (In Russ.)

Сведения об авторе / Information about the author

Артемова Анастасия Николаевна – кандидат юридических наук, научный сотрудник Института философии и права Сибирского отделения Российской академии наук, г. Новосибирск, ул. Николаева, 8, e-mail: artemova-an-1991@yandex.ru.

Статья поступила в редакцию: 15.08.2025

После доработки: 08.09.2025

Принята к публикации: 20.09.2025

Artemova Anastasiia – Candidate of Juridical Sciences, Researcher of the Institute of Philosophy and Law of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Nikolaeva Str., 8, e-mail: artemova-an-1991@yandex.ru.

The paper was submitted: 15.08.2025

Received after reworking: 08.09.2025

Accepted for publication: 20.09.2025