УДК 004.89

ИНСТРУМЕНТЫ ОБЕСПЕЧЕНИЯ ЛИЦЕНЗИРУЕМОСТИ И НОРМАТИВНОЙ СОВМЕСТИМОСТИ В ML-БИБЛИОТЕКАХ

Смирнов Д.А., студент гр. Б02-307, I курс
Московский физико-технический институт, г. Москва

С переходом машинного обучения из исследовательских лабораторий в сферы государственного управления, здравоохранения, финансов и промышленности на первый план выходят не только вопросы точности и эффективности алгоритмов, но и проблемы правового соответствия, лицензируемости и нормативной совместимости. Использование ML-библиотек в регулируемых секторах требует строгого соблюдения юридических норм, соответствия отраслевым стандартам, открытости происхождения компонентов и возможности последующего лицензирования решений. Всё это делает задачу обеспечения лицензируемости и нормативной совместимости не факультативной, а фундаментальной частью архитектуры доверенных библиотек машинного обучения. Настоящая статья посвящена всестороннему анализу инструментов и подходов, направленных на достижение юридической прозрачности, соблюдение нормативов и возможность повторного использования компонентов на легитимных основаниях [1].

Под лицензируемостью в контексте ML-библиотек понимается не только корректное соблюдение условий лицензий (например, Apache 2.0, MIT, GPL), но и возможность воспроизвести цепочку происхождения всех зависимостей, моделей, данных и скриптов, входящих в состав программного продукта. Это особенно важно в условиях, когда система включает десятки сторонних модулей с разными условиями использования. Наличие инструментов отслеживания лицензионного статуса компонентов и управления лицензиями (license compliance management) становится критически необходимым. Для этого применяются специализированные средства, такие как FOSSA, ClearlyDefined, OpenChain, ScanCode и SPDX (Software Package Data Exchange), которые позволяют автоматически анализировать зависимости, выявлять лицензии, проверять совместимость и документировать результаты в машиночитаемом формате.

Важнейшим направлением является инвентаризация и паспортизация программных компонентов. Для каждого элемента библиотеки должно быть зафиксировано: происхождение (включая авторов и источник), версия, тип

лицензии, ограничения на использование (например, запрет на коммерческое применение), а также хэш-суммы или другие идентификаторы для верификации подлинности. Такие паспорта позволяют организациям формировать доверенные реестры компонентов и использовать только проверенные модули в продуктивной среде. В более сложных случаях возможно применение белых и чёрных списков – curated lists – поддерживаемых юридическими или отраслевыми организациями [2].

Особое внимание уделяется юридической совместимости лицензий, особенно в случае комбинирования компонентов с различными условиями. Нарушение условий (например, смешение GPL с закрытым кодом) может привести к юридическим санкциям, отзывам продуктов или даже судебным искам. Инструменты автоматического анализа (например, license scanners) помогают выявить потенциальные конфликты и предложить альтернативы. Всё чаще в корпоративной практике применяются policy-as-code системы, где правила лицензирования кодируются в виде политик, автоматически применяемых к новым компонентам при их включении в проект.

Параллельно с вопросами лицензий стоит проблема нормативной совместимости, особенно для библиотек, применяемых в отраслях с высокими требованиями к безопасности, объяснимости и этичности систем. В ЕС, например, вступление в силу AI Act потребует от разработчиков машинного обучения демонстрировать соответствие требованиям прозрачности, устойчивости и надёжности. Аналогично, в здравоохранении системы должны соответствовать стандартам HIPAA, в промышленности – IEC 61508, в автомобилестроении – ISO 26262, а в финансовом секторе – стандартам Basel III и GDPR. Это означает, что библиотеки ML, предназначенные для доверенного применения, должны не только обеспечивать техническую функциональность, но и включать встроенные средства документирования и проверки соответствия нормативам [3].

Для этого разрабатываются метаданные соответствия, где для каждого модуля указывается его уровень готовности к сертификации, включённые механизмы защиты (например, дифференциальная приватность, TEE), а также логика, соответствующая регулируемым сценариям (например, запрет на автоматическое принятие решений без участия человека). Такие метаданные оформляются в виде деклараций соответствия (compliance declarations) или аннотаций, прикрепляемых к каждому релизу программного пакета. Некоторые фреймворки начинают поддерживать автоматическую генерацию таких документов с помощью CI/CD-конвейеров, что позволяет системам отслеживать нормативный статус в режиме реального времени.

Кроме того, важную роль играют репозитории доверенных компонентов, включающие только те модули, которые прошли юридическую и нормативную

экспертизу. Организации и отрасли могут формировать свои внутренние или открытые каталоги сертифицированных библиотек, что повышает уровень доверия и снижает риски. Интеграция таких репозиториев с ML-фреймворками позволяет автоматически запрещать или помечать модули, не прошедшие проверку, и тем самым реализовывать доверенные цепочки поставок ПО (trusted software supply chains) [4].

Наконец, обеспечение лицензируемости и нормативной совместимости требует определённого уровня образования и инструментальной поддержки для разработчиков. Создание библиотек, которые «по умолчанию соответствуют требованиям» (compliance by design), возможно только при наличии стандартизированных шаблонов, генераторов лицензий, валидаторов политик и механизмов обратной связи. Это направление активно развивается, включая инициативы по сертификации библиотек и фреймворков, стандартизации описаний (например, с использованием SPDX и OpenSSF) и формированию обучающих платформ по юридическим аспектам разработки в области ИИ.

В заключение следует подчеркнуть, что обеспечение лицензируемости и нормативной совместимости в ML-библиотеках – это не вспомогательная задача, а необходимое условие для превращения моделей машинного обучения в легитимные, подотчётные и этически допустимые программные решения. В условиях растущего регулирования и общественного внимания к ИИ-продуктам, формирование инфраструктуры, способной автоматически контролировать лицензионную чистоту и нормативное соответствие, становится одним из центральных направлений в развитии доверенных систем. Дальнейшие усилия в этой области должны быть сосредоточены на автоматизации, стандартизации и создании глобальных экосистем доверенного, правового и регулируемого машинного обучения [5].

Список литературы:

1. Open Source Initiative. The Open Source Definition [Электронный ресурс]. – URL: https://opensource.org/osd (дата обращения: 09.06.2025).

2. Apache Software Foundation. Apache License, Version 2.0 [Электронный ресурс]. – URL: https://www.apache.org/licenses/LICENSE-2.0 (дата обращения: 09.06.2025).

3. Free Software Foundation. GNU General Public License, Version 3 [Электронный ресурс]. – URL: https://www.gnu.org/licenses/gpl-3.0.html (дата обращения: 09.06.2025).

4. Linux Foundation. SPDX Specification, Version 2.3 [Электронный ресурс]. – URL: https://spdx.github.io/spdx-spec/v2.3/ (дата обращения: 09.06.2025).

5. Creative Commons. Creative Commons Attribution 4.0 International Public License [Электронный ресурс]. – URL: https://creativecommons.org/licenses/by/4.0/legalcode (дата обращения: 09.06.2025).