



Возможность применения моделей искусственного интеллекта в рамках рутинных процессов фармаконадзора

Широбоков Я. Е., Молчанова Ю. А., Чумак Е. П., Скрипкин А. Ю.

АО «Р-Фарм», Москва, Российская Федерация

Аннотация

Введение. Искусственный интеллект (ИИ) прошёл стремительный путь развития в фармаконадзоре (ФН): от экспериментального применения до рассмотрения в качестве одного из ключевых инструментов повседневной практики. Относительно простые модели искусственного интеллекта (ИИ-модель), включая статистические методы выявления сигналов безопасности, используются в ФН на протяжении десятилетий, тогда как последние достижения в области обработки естественного языка (Natural Language Processing; NLP) существенно расширили спектр возможного применения.

Цель. Настоящая статья посвящена критическому анализу потенциала NLP-систем для оптимизации рутинных задач ФН с учётом требований к защите данных. Отдельно рассматривается применение ИИ-моделей семантического поиска на основе иных архитектурных подходов, в частности embedding-моделей и retrieval-augmented generation.

Основные положения. Выделяются процессы, не предполагающие работу с персональными данными и допускающие использование открытых ИИ-решений (поиск и систематизация научной литературы, генерация резюме публикаций), а также процессы, связанные с обработкой конфиденциальной информации (например, извлечение данных из индивидуальных сообщений о нежелательных реакциях, автоматическое формирование описаний клинических случаев, анализ соотношения «польза-риск», проверка соответствия регуляторным требованиям и стандартам отчётности), для которых требуется применение корпоративных ИИ-систем, развернутых в защищённой инфраструктуре. В статье также рассматриваются ограничения и риски, практические аспекты внедрения, а также вопросы обеспечения надёжности, воспроизводимости и прозрачности используемых решений.

Заключение. ИИ-модели и, в частности, NLP-модели, обладают значительным потенциалом для внедрения в рутинные процессы ФН, т.к. способны существенно сократить временные затраты на выполнение рутинных задач. Вместе с тем успешная интеграция ИИ-моделей невозможна без системного подхода к управлению сопутствующими рисками. При соблюдении этих условий ИИ способен стать полноценным инструментом повышения эффективности процессов ФН.

Ключевые слова: фармаконадзор; искусственный интеллект; NLP; LLM; CIOMS; рутинные процессы

Для цитирования: Широбоков Я. Е., Молчанова Ю. А., Чумак Е. П., Скрипкин А. Ю. Возможность применения моделей искусственного интеллекта в рамках рутинных процессов фармаконадзора. *Реальная клиническая практика: данные и доказательства*. 2026;6(2):56-67. <https://doi.org/10.37489/2782-3784-myrd-104>. EDN: ZBXLQV.

Поступила: 18.05.2026. **В доработанном виде:** 20.06.2026. **Принята к печати:** 26.06.2026. **Опубликована:** 30.06.2026.

Possibility of using artificial intelligence models in routine pharmacovigilance processes

Yaroslav E. Shirobokov, Yuliya A. Molchanova, Eugenia P. Chumak, Aleksei Yu. Skripkin

JSC R-Pharm, Moscow, Russian Federation

Abstract

Introduction. Artificial intelligence (AI) has undergone rapid development in pharmacovigilance (PV), evolving from experimental application to being considered a key tool in daily practice. Relatively simple AI models, including statistical signal detection methods, have been used in PV for decades, while recent advances in Natural Language Processing (NLP) have significantly expanded the scope of potential applications.

Objective. This article provides a critical analysis of the potential of NLP systems to optimize routine PV tasks, taking into account data protection requirements. The application of semantic search AI models based on alternative architectural approaches, specifically embedding models and retrieval-augmented generation (RAG), is examined separately.

Main points. The authors distinguish between processes that do not involve personal data and allow the use of open AI solutions (searching and systematizing scientific literature, generating publication summaries), and processes involving the handling of confidential information (e. g., data extraction from Individual Case Safety Reports (ICSRs), automated generation of clinical case descriptions, benefit-risk analysis, and compliance with regulatory requirements and reporting standards), which require the use of corporate AI systems deployed within a secure infrastructure. The article also discusses limitations, risks, practical implementation aspects, as well as issues of ensuring the reliability, reproducibility, and transparency of the solutions used.

Conclusion. AI models, particularly NLP models, have significant potential for integration into routine PV processes. However, successful integration of AI models is impossible without a systematic approach to managing associated risks. Subject to these conditions, AI can become a robust tool for increasing the efficiency of PV processes.

Keywords: pharmacovigilance; artificial intelligence; NLP; LLM; CIOMS; routine processes

For citation: Shirobokov YaE, Molchanova YuA, Chumak EP, Skripkin AYU. Possibility of using artificial intelligence models in routine pharmacovigilance processes. *Real-World Data & Evidence*. 2026;6(2):56-67. <https://doi.org/10.37489/2782-3784-myrd-104>. EDN: ZBXLQV.

Received: 18.05.2026. **Revision received:** 20.06.2026. **Accepted:** 26.06.2026. **Published:** 30.06.2026.

Введение / Introduction

Современный этап развития общества характеризуется стремительным внедрением цифровых технологий, среди которых особое место занимает искусственный интеллект (ИИ) [1]. Особую роль в развитии современных интеллектуальных систем играют нейронные сети — математические модели, функционирующие по принципу биологических нейронных систем человека [2]. Нейронные сети представляют собой совокупность взаимосвязанных искусственных нейронов, организованных в слои, где каждый нейрон принимает входные данные, обрабатывает их и передает результат следующему слою. В процессе обучения нейронная сеть корректирует весовые коэффициенты связей между нейронами, увеличивая вес правильных ответов и уменьшая вес ошибочных [3]. Современные нейросетевые технологии позволяют выполнять обработку больших массивов данных, выявлять скрытые закономерности, прогнозировать результаты и автоматизировать процессы принятия решений. Благодаря способности к самообучению и адаптации нейросети получили широкое распространение в медицине, экономике, промышленности, образовании, аналитике данных и информационных технологиях.

В настоящее время существует множество архитектур нейронных сетей, различающихся принципами организации и обработки информации. В числе примеров архитектур нейронных сетей можно выделить нейронные сети прямого распространения (Feed Forward Neural Network; FFNN), рекуррентные нейронные сети (Recurrent Neural Network; RNN), сети долгой краткосрочной памяти (Long Short-Term Memory; LSTM) и свёрточные нейронные сети (Convolutional Neural Network; CNN) [4].

Отдельным и наиболее современным направлением развития нейронных сетей является

архитектура, основанная на механизме внимания, — в частности, трансформер (Transformer). Трансформер — это архитектура нейронной сети, которая наиболее эффективна для обработки последовательностей данных. Трансформеры, как и рекуррентные нейронные сети, также предназначены для обработки последовательностей, таких как текст на естественном языке [5, 6].

Одной из разновидностей ИИ являются большие языковые модели (Large Language Models; LLM), основанные на архитектуре трансформеров и обученные на огромных массивах текстовых данных для выполнения широкого спектра задач, связанных с обработкой естественного языка (Natural Language Processing; NLP), таких как генерация текста, перевод, суммаризация, вопросно-ответные системы, кодирование и другие [7].

Основной задачей NLP является обеспечение эффективного взаимодействия между человеком и компьютерной системой посредством естественного языка. Современные NLP-модели применяются в интеллектуальных поисковых системах, голосовых помощниках, чат-ботах, системах автоматического перевода, анализе тональности текста, обработке документов и генерации текстовой информации [8]. Интеллектуальный анализ текста включает ряд методов, используемых для характеристики и преобразования текста. В рамках интеллектуального анализа текста NLP представляет собой набор синтаксических и/или семантических алгоритмов обработки, основанных на правилах или статистических данных, которые могут использоваться для сегментации, извлечения или анализа текстовых данных. Различие подходов состоит в том, что интеллектуальный анализ текста использует сами слова в качестве единицы анализа

(например, частота, наличие или отсутствие конкретных интересующих слов), в то время как методы NLP используют базовые метаданные, включая содержание и шаблоны фраз. И NLP, и интеллектуальный анализ текста используются в ряде онлайн-областей, связанных со здоровьем, таких как психическое здоровье, онкология и инфекционные заболевания [9].

В настоящее время NLP-модели активно реализуются как в отечественных, так и в зарубежных интеллектуальных системах. Среди российских разработок можно выделить такие платформы как GigaChat, Алиса, Gureev Pro AI. Среди зарубежных NLP-платформ наибольшее распространение получили Claude AI, ChatGPT, Google Gemini, DeepSeek, Kimi, Qwen и Perplexity AI. Данные системы основаны на архитектуре Transformer и способны выполнять глубокий анализ информации, генерировать тексты, осуществлять интеллектуальный поиск, обрабатывать научные материалы, создавать программный код и автоматизировать аналитическую деятельность [6].

Развитие технологий ИИ открывает новые возможности для автоматизации рутинных процессов фармаконадзора, однако их внедрение требует оценки эффективности, надёжности и соответствия требованиям защиты данных. Актуальность данной проблемы обусловлена ростом объёма информации о безопасности лекарственных средств и необходимостью повышения качества и скорости её обработки. Исследование возможностей применения NLP-систем и других ИИ-моделей имеет важное значение для совершенствования процессов мониторинга безопасности лекарственных препаратов и дальнейшего развития цифровых технологий в фармацевтической отрасли.

Применение ИИ-моделей для оптимизации рутинных процессов фармаконадзора и преимущества их внедрения / Application of AI models for optimizing routine pharmacovigilance processes and the benefits of their implementation

При оценке потенциала внедрения ИИ-моделей в процессы ФН необходимо принимать во внимание ряд существенных аспектов, подробно изложенных в отчёте рабочей группы XIV Совета международных организаций медицинских наук (CIOMS) «Искусственный интеллект в фармаконадзоре» [10].

Одним из фундаментальных аспектов среди них является защита конфиденциальности данных. В рамках своей профессиональной деятельности специалисты по ФН работают с персональными данными пациентов, а действующее законодательство, нормативные акты и руководящие документы предусматривают комплекс мер, направленный на обеспечение конфиденциальности, анонимности,

автономии и контроля в отношении чувствительных и потенциально идентифицируемых медицинских данных. Помимо этого, значительная часть деятельности в сфере ФН связана с внутренними документами компании, которые в соответствии с соглашениями о конфиденциальности не подлежат разглашению третьим сторонам.

Совокупность указанных ограничений исключает возможность использования общедоступных решений, таких как Claude, ChatGPT, Google Gemini и других, несмотря на то что данные модели представляют собой готовые к применению инструменты с широким спектром функциональных возможностей, не требующие длительного обучения. Вместе с тем ряд процессов в ФН предполагает работу с открытыми данными, что допускает применение перечисленных моделей. Указанные процессы рассматриваются ниже.

Поиск и систематизация научной литературы / Scientific literature search and systematization

Поиск научной литературы — один из процессов, охватывающий практически все направления ФН: оценку причинно-следственной связи, анализ сигналов, составление периодических отчётов по безопасности (ПООБ) и планов управления рисками (ПУР), подготовку обучающих материалов в рамках дополнительных мер минимизации рисков, проведение оценки соотношения «польза-риск», подготовку ответов на запросы регуляторных органов и ряд других задач.

Необходимо подчеркнуть, что в данной главе не рассматривается применение ИИ для мониторинга научной медицинской литературы на еженедельной основе или за определённый временной период. Речь идёт о возможности тематического поиска литературы по заранее заданным критериям.

Также данное направление является единственным в настоящей статье, где рассматривается применение не NLP-моделей, а моделей семантического поиска.

Процесс поиска нужной литературы зачастую осуществляется вручную. Специалист обращается к библиографическим базам данных (PubMed, Google Scholar, КиберЛенинка, eLIBRARY), формирует запрос с применением ключевых слов, фильтров, логических операторов (И, ИЛИ, НЕ), после чего последовательно просматривает заголовки и аннотации статей для оценки их релевантности. Данный процесс достаточно трудоёмок. Можно столкнуться как с необходимостью просмотра значительного массива публикаций, большинство из которых не отвечают критериям отбора, так и с малым числом публикаций или их полным отсутствием. В результате приходится расширять запрос, убирать фильтры или обобщать критерии, что

неизбежно влечёт за собой дополнительный объём работы, который не всегда приводит к нужному результату.

Применение ИИ-моделей способно существенно оптимизировать данный процесс, прежде всего за счёт исключения этапа сплошного просмотра всех предложенных публикаций. По сравнению с ручным поиском в библиографических базах данных применение ИИ-моделей обладают рядом значимых преимуществ:

- более высокая точность поиска. При грамотно составленном запросе ИИ-модель, как правило, выдаёт релевантные статьи с их кратким описанием. Это позволяет оперативно оценить соответствие публикации заданным критериям;
- расширенные возможности поиска по критериям, недоступным в стандартных библиографических базах. Например, поиск метаанализов или систематических обзоров с наличием показателей относительного или абсолютного риска, ретроспективных исследований с анализом факторов риска, или описаний клинических случаев с данными о времени развития нежелательной реакции от момента начала применения препарата;
- информирование об отсутствии релевантных публикаций. В отличие от ручного поиска, при котором специалист тратит время, прежде чем убедиться в отсутствии статей по заданной теме, ИИ-модель сообщает об этом незамедлительно. Это позволяет сразу скорректировать критерии поиска или принять решение об отсутствии статей по рассматриваемому вопросу;
- существенное сокращение времени. Благодаря повышенной точности и расширенным критериям поиска значительно снижаются временные затраты на получение необходимых данных.

Таким образом, применение ИИ-моделей для поиска и систематизации научной литературы позволяет значительно повысить скорость и эффективность одного из наиболее трудоёмких вспомогательных процессов в ФН.

Извлечение данных с помощью больших языковых моделей, генерация резюме публикаций / Data Extraction Using Large Language Models and Generation of Publication Summaries

Следующим этапом после нахождения релевантной публикации является её прочтение с целью извлечения необходимых данных. Опытные специалисты, как правило, не читают статью целиком, а целенаправленно ищут нужную информацию по ключевым словам или обращаются непосредственно к соответствующему разделу. Тем не менее даже такой подход требует времени и не исключает риска пропуска важных данных, особенно если искомая

информация содержится в объёмных таблицах или графиках, или специалист работает в условиях высокой нагрузки.

Применение LLM позволяет существенно ускорить этот процесс. Достаточно сформулировать запрос с указанием необходимых данных, после чего модель либо представит резюме запрашиваемой информации, либо сообщит об её отсутствии в статье. Дополнительным преимуществом данного подхода является простота верификации ответа, так как любой фрагмент, сгенерированный моделью, можно проверить, найдя соответствующий текст в исходной публикации. Практическим примером применения данного сценария может служить извлечение данных для характеристики рисков в разделе 16.4 ПООБ или модуле CVII ПУР (например, факторов риска, времени развития нежелательной реакции от начала применения препарата, специфических клинических проявлений, исходов и мер по обратимости).

Другим сценарием является подготовка резюме публикации, например, для разделов 11, 17.1, 17.2, 18.1 ПООБ или модуля CI ПУР. В данном случае достаточно указать, какие аспекты статьи должны быть отражены в резюме. Следует учитывать, что стиль изложения, генерируемый моделью при первых итерациях, может не в полной мере соответствовать требованиям конкретного документа. Для достижения приемлемого результата необходима предварительная калибровка модели. При работе с открытыми решениями, в большинстве случаев, достаточно вручную отредактировать несколько текстов, загрузить их в качестве образцов и задать инструкцию придерживаться заданного стиля. После такой калибровки модель становится готовым инструментом, способным значительно сократить временные затраты на подготовку документов.

Применение LLM в описанном ключе возможно и для работы с внутренними документами компании, например, для извлечения данных о нежелательных явлениях, зафиксированных в ходе клинического исследования, из соответствующего отчёта. Однако в этом случае использование открытых моделей недопустимо и требуются внутренние, обученные на корпоративных данных решения с налаженной системой периодического контроля качества. Данный аспект будет подробнее рассмотрен ниже.

Подготовка и обработка индивидуальных сообщений о нежелательных реакциях / Preparation and Processing of Individual Case Safety Reports (ICSRs)

Число ИСНР, поступающих в отделы ФН фармацевтических компаний, неуклонно растёт. При этом качество и объём содержащихся в них данных существенно варьируются: от кратких сообщений

с минимальным набором сведений (демографические характеристики пациента и описание нежелательной реакции) до обширных неструктурированных текстов или копий документов в формате pdf или jpg, включающих сложные лабораторные показатели и клинически значимую информацию, которую специалист может не заметить при ручной обработке. Примерами такой информации служат несоответствие фактического показателя к применению препарата зарегистрированному показанию (например, его off-label назначение при почечноклеточном раке вследствие ошибочной интерпретации, вызванной визуальным сходством терминов с зарегистрированным показанием, — лечение печёчноклеточного рака) или ошибки в расчёте дозы относительно массы тела пациента и др. Дополнительную нагрузку создают периоды резкого увеличения потока ИСНР (как это наблюдалось в период пандемии коронавирусной инфекции), которые при ограниченном штате сотрудников неизбежно ведут к снижению качества обработки сообщений и риску пропуска клинически важных сведений. Вместе с тем фармацевтические компании сохраняют обязанность своевременно реагировать на сообщения о нежелательных реакциях в целях защиты общественного здоровья и оперативного выявления и устранения проблем по безопасности.

Ранее для решения этих задач применялся подход на основе распознавания именованных сущностей (Named Entity Recognition; NER) — технология, позволяющая выявлять и классифицировать ключевые элементы информации в неструктурированных текстах (наименования нежелательных реакций, квалификация репортера, страна, специфические термины) с использованием заранее заданной разметки. Однако на практике данный подход не обеспечил стабильного улучшения процессов обработки ИСНР [10].

Применение LLM открывает новые возможности для ускорения обработки ИСНР при сохранении качества. Хорошо обученная LLM способна извлекать из исходных данных (например, E2B-файлов, форм CIOMS, публикаций в научной литературе) заранее определённый перечень сведений: 4 минимальных критерия валидности ИСНР, медицинский анамнез, дату начала применения подозреваемого препарата, время развития и разрешения нежелательной реакции, исход и сопутствующую терапию. На основе извлечённых данных модель может проводить анализ случая, оценку причинно-следственной связи и определять недостающую информацию, которую необходимо запросить для полноценной оценки. Следующим шагом LLM формирует структурированное резюме случая в заранее установленном формате и стиле изложения (например, по логике *anamnesis morbi*).

Рабочая группа XIV CIOMS в своем отчёте [10] уже описала практический сценарий применения LLM для извлечения данных при обработке ИСНР. В нем представлены результаты экспериментального исследования, в рамках которого был выстроен полуавтоматический процесс обработки ИСНР, состоящий из трёх этапов. На первом этапе входные документы различных форматов приводились к единому виду. Затем модель GPT-4 извлекала данные с помощью структурированного JSON-шаблона, содержащего подсказки о вероятном расположении информации в документе. На третьем этапе текстовые ответы модели автоматически сопоставлялись с предопределёнными значениями полей. Для оценки эффективности был разработан графический интерфейс, позволявший сравнить полностью ручную обработку с процессом, дополненным ИИ. Четыре опытных специалиста были случайным образом распределены между двумя вариантами работы, а система фиксировала чистое время обработки.

Согласно полученным результатам, точность извлечения составила 85–100 % для документов по клиническим исследованиям, 60–100 % для сообщений из программ поддержки пациентов и 67–100 % для случаев из медицинской литературы. Применение LLM сократило затраты времени на 39 %, что позволило сэкономить в среднем около 20 минут на каждое извещение. Из примерно 69 точек данных, извлечённых моделью, ручной корректировки требовали лишь 2,4 [10].

Таким образом, такой подход позволяет существенно сократить время обработки ИСНР, снизить риск пропуска клинически значимой информации и максимально унифицировать стиль описания случаев.

Подготовка справочных, информационных материалов и адаптация текстов / Preparation of Reference and Information Materials and Text Adaptation

Важным направлением ФН является подготовка информационных материалов для специалистов здравоохранения (например, руководств для врачей) и пациентов (например, памяток) в рамках дополнительных мер минимизации рисков.

Согласно требованиям надлежающей практики ФН [11], образовательные инструменты должны быть адаптированы к потребностям целевой аудитории. На практике это означает необходимость обеспечить понятность текста как для специалистов здравоохранения, так и для пациентов, не имеющих медицинского образования. При этом адаптация профессионального медицинского текста, содержащего узкоспециализированную терминологию, к формату, доступному для пациента представляет

некоторые затруднения даже для опытного специалиста и требует временных затрат. В случае препаратов с инновационным механизмом действия (например, средств генной терапии) подготовка доступного обучающего материала может вызывать затруднения и для аудитории врачей.

Применение LLM способно существенно ускорить этот процесс. Модель может оперативно формулировать сложные описания, адаптируя язык и стиль изложения под конкретную целевую аудиторию — от специалиста здравоохранения до пациента. По аналогии с обработкой ИСНР, описанной выше, LLM способна формировать структурированный текст на основе имеющихся исходных материалов (например, общая характеристика лекарственного препарата, научные публикации, иные информационные источники) в соответствии с заранее заданной структурой или шаблоном. Данный подход применим и при подготовке разделов по безопасности листка-вкладыша для пациента. LLM позволяет адаптировать текст из общей характеристики лекарственного препарата, упрощая медицинский нарратив при полном сохранении его смысла.

Таким образом, применение LLM при подготовке информационных материалов позволяет существенно сократить временные затраты на создание и адаптацию текстов для различных целевых аудиторий. Особую ценность данный подход представляет в работе со сложными терапевтическими областями, где адаптация профессионального медицинского текста для пациентов является трудоёмкой задачей.

Анализ соотношения «польза-риск» / Benefit-Risk Analysis

Одной из ключевых целей ФН является оценка того, превышает ли польза от применения лекарственного препарата его потенциальные риски, по результатам анализа накопленных данных о безопасности. Внедрение NLP в процесс оценки соотношения «польза-риск» представляет собой перспективное направление развития ФН.

На сегодняшний день разработано множество методов оценки потенциальной пользы и рисков лекарственной терапии, которые классифицируются на количественные, полуколичественные и качественные. Тем не менее на практике преобладают качественные методы, среди которых наибольшее распространение получили PrOACT-URL, Benefit-Risk Framework, модель BRAT (Benefit-Risk Action Team) и ряд других [12]. Несмотря на методологические различия между ними, все перечисленные подходы можно интегрировать с NLP-системами.

Вместе с тем, из всех рассматриваемых в настоящей статье направлений применения ИИ данное направление является наиболее трудоёмким с точки

зрения реализации. Качественные методики, как правило, предполагают ответы на определённый набор вопросов. Например:

- какое заболевание подлежит лечению и какова степень его тяжести;
- в чём выражается ожидаемая польза (излечение, продление жизни, купирование симптомов);
- какие риски сопряжены с применением препарата, какова их частота и управляемость;
- существуют ли альтернативные варианты терапии и каковы их сравнительные преимущества и недостатки.

Для формирования ответа на каждый из этих вопросов модели необходимы специфические обучающие данные: общая характеристика лекарственного препарата, отчёты о клинических исследованиях, эпидемиологические данные, клинические рекомендации и другие источники. Таким образом, разработка подобной модели требует значительных временных затрат на сбор данных, обучение и последующую валидацию качества.

Тем не менее, внедрение такой модели позволяет автоматизировать рутинную часть работы, а именно анализ и систематизацию больших массивов данных с последующим их описанием, оставляя специалисту самый ответственный этап — итоговое взвешивание и принятие решения. Такой подход обеспечивает прозрачность оценки для регуляторных органов и способствует повышению качества, объективности, доказательности, воспроизводимости и достоверности проводимой оценки, снижая влияние человеческого фактора, когда специалист может непреднамеренно акцентировать внимание на данных, подтверждающих его мнение о препарате. При этом принципиально важно, что ИИ не замещает экспертное суждение специалиста, а усиливает его, концентрируя человеческий ресурс на этапах, требующих профессиональной и этической ответственности.

Пример применения ИИ-модели для схожего сценария представлен в отчёте CIOMS [10]. Рабочая группа исследовала применимость ИИ-моделей для автоматизации, стандартизации и поддержки принятия экспертных решений в области оценки причинно-следственной связи. Для обучения и тестирования моделей использовался обширный гетерогенный набор данных, объединявший спонтанные сообщения, данные клинических исследований и медицинскую литературу. Набор включал демографические характеристики пациентов, анамнез, сопутствующие заболевания, широкий спектр лекарственных препаратов и вакцин, а также разнообразные типы нежелательных явлений, закодированных по терминологии MedDRA. Архитектура модели предусматривала извлечение признаков из

структурированных полей и текстовых нарративов с применением NLP и LLM, расчёт суррогатных метрик, в частности, по шкале Наранжо или критериям ВОЗ, и финальную классификацию степени причинно-следственной связи от «сомнительной» до «определённой».

Данная модель показала высокую воспроизводимость результатов, существенно снижая субъективный фактор. ИИ-модель эффективно учитывала скрытые кофакторы, такие как возраст, сопутствующую терапию, которые нередко упускаются при ручном анализе. ИИ-модель успешно классифицировала случаи по уровню неопределённости: очевидные и хорошо задокументированные реакции обрабатывались автоматически, тогда как сложные и пограничные направлялись на углублённый консиллум к экспертам [10].

Несмотря на то, что данный пример относится к оценке причинно-следственной связи, а не к анализу соотношения «польза-риск» напрямую, он наглядно демонстрирует принципиальную применимость аналогичного подхода и в этой области. Логика процесса во многом схожа: обучение модели на большом массиве разнородных данных с последующим заполнением заранее определённых полей на их основе. Отличие состоит лишь в том, что в случае анализа соотношения «польза-риск» эти поля определяются не шкалой Наранжо или критериями ВОЗ, а, например, структурой методологии PrOACT-URL.

Проверка соответствия регуляторным требованиям и стандартам отчётности / Regulatory Compliance and Reporting Standards Verification

Все документы в сфере ФН должны соответствовать установленным требованиям и следовать определённой структуре, включая обязательные разделы, регламентированные документами соответствующих регуляторных органов и отраслевых стандартов (ЕЭК, CIOMS, FDA, EMA, ICH).

Несмотря на то, что основные требования к структуре документов унифицированы, отдельные аспекты могут варьироваться, что создаёт дополнительные сложности для фармацевтических компаний, осуществляющих деятельность в нескольких странах с разными регуляторными органами. Необходимость детального изучения каждого регуляторного документа и шаблона с целью выявления таких различий является трудоёмкой и времязатратной задачей, не исключающей риска пропуска важных требований и, как следствие, отклонений и повторных подач документов.

Одним из инструментов решения данной проблемы может служить применение NLP-моделей в двух основных сценариях.

Первый сценарий — проверка соответствия формальным требованиям к структуре документов. NLP-модели с шаблонным распознаванием или специализированные чек-листы, интегрированные с LLM, способны верифицировать наличие всех обязательных разделов и выявлять несоответствия форматам CIOMS, FDA и других регуляторных органов.

Второй сценарий — проверка соответствия содержания документа. NLP-модель может анализировать текст документа и сопоставлять его с требованиями руководящих документов регуляторов. Например, проверять, соответствует ли структура ПООБ рекомендациям ICH E2C (R2), отражены ли требования модуля IX Надлежащей практики фармаконадзора в описании сигнала безопасности, корректно ли интерпретированы серьёзность нежелательной реакции и причинно-следственная связь. Особую практическую ценность данный подход представляет в ситуациях, когда одно и то же ИСНР необходимо направить одновременно в Росздравнадзор, FDA и EMA, каждый из которых предъявляет собственные требования к языку, формату и степени детализации.

Внедрение NLP-моделей в процессы проверки документации позволяет специалисту по ФН сосредоточиться на содержательной и медицинской составляющей документа, не отвлекаясь на верификацию соответствия формальным требованиям. Ожидается также снижение числа запросов со стороны регуляторных органов, обусловленных формальными несоответствиями. В случае же поступления подобных запросов NLP-модели могут дополнительно применяться для анализа и систематизации замечаний регулятора, формируя тем самым расширенную базу требований, которая будет учитываться при последующей проверке документов.

Основные ограничения и риски ИИ-моделей / Key Limitations and Risks of AI Models

В предыдущих главах рассмотрены преимущества внедрения ИИ-моделей в практическую деятельность специалиста по ФН, а именно сокращение времени выполнения рутинных задач, повышение стандартизации и качества документов, возможность сосредоточиться на экспертной аналитической работе, а также снижение влияния человеческого фактора при обработке больших объёмов данных. Вместе с тем применение ИИ-моделей сопряжено с рядом существенных ограничений и рисков, которые рассматриваются в данном разделе.

Риск ошибок и недостоверной генерации («галлюцинации»). Несмотря на постепенное повышение качества современных LLM, они по-прежнему

подвержены так называемым «галлюцинациям» — генерации правдоподобных, но фактически недостоверных ответов, содержащих сфабрикованные факты, ложные ссылки или логические ошибки. В контексте ФН это может повлечь серьезные последствия, среди которых: неверная оценка серьезности случая и, как следствие, несвоевременная подача ИСНР в регуляторный орган, выявление ложных сигналов или наоборот пропуск существующих, некорректная оценка соотношения «польза-риск», неточные формулировки в памятке для пациента или листке-вкладыше, вводящие пациента в заблуждение. Поскольку точность информации в ФН непосредственно влияет на безопасность пациентов и своевременность выявления проблем безопасности, риск недостоверной генерации должен быть в обязательном порядке учтен при планировании применения LLM [10].

Аналогичный риск существует и при использовании моделей семантического поиска для систематизации литературы. При отсутствии публикаций по заданным критериям модель может сгенерировать несуществующую статью с правдоподобно указанными авторами, названием журнала и даже идентификатором DOI.

Недетерминированность. Генеративные модели не детерминированы. То есть они выдают разные ответы на один и тот же запрос без изменения самой модели. Это затрудняет проверку стабильности системы [10].

Зависимость от качества и количества данных. Качество работы ИИ-моделей зависит от качества обучающих данных. Однако данные ФН часто характеризуются несогласованностью, неполнотой, ошибками кодирования и высокой вариативностью. Также ИИ-модели могут показывать неудовлетворительные результаты для отдельных групп населения (разделение по расе, гендеру, возрасту или социально-экономическим факторам) или редких нежелательных реакций, если эти данные были мало представлены в обучающей выборке. Всё это снижает точность ИИ-моделей [10].

Перенос и усиление предвзятости человека. Поскольку ИИ-модели обучаются на данных, созданных людьми, они будут перенимать и, возможно, даже усиливать скрытую или явную предвзятость и дискриминационные паттерны человека [10].

Риск утечки конфиденциальных данных. Как отмечалось ранее, использование открытых ИИ-моделей при работе с персональными данными пациентов и внутренней документацией компании категорически недопустимо. Вместе с тем даже при разработке собственной ИИ-модели, функционирующей в рамках корпоративной инфраструктуры, полностью исключить риск утечки данных не представляется возможным [10].

Описанные ограничения и риски носят критически важный характер при планировании внедрения ИИ-моделей в рутинные процессы ФН и обуславливают необходимость валидации, систематической проверки результатов («фактчекинг»), контроля качества обучающих и последующих генерируемых данных, а также реализации технических мер защиты от несанкционированного доступа и утечек информации. Указанные меры подробно рассматриваются в следующем разделе.

Практические аспекты внедрения / Practical Aspects of Implementation

Для полноценного внедрения ИИ-моделей в деятельность специалиста по ФН необходимо обеспечить достоверность и надёжность принимаемых ими решений. Достоверность означает, что модель достигает поставленной цели в пределах допустимых параметров [10]. Для этого требуется предварительное определение приемлемых показателей производительности, подбора репрезентативных данных для обучения и тестирования, оценка эффективности модели в реалистичных условиях и её интеграции в действующую систему обеспечения качества. Надёжность, в свою очередь, означает стабильное достижение поставленных целей с учётом вариативности входных данных. В своем отчёте CIOMS выделяет 5 ключевых аспектов, которые необходимо принимать во внимание при внедрении ИИ-модели [10].

Спецификация задачи предполагает чёткое определение цели и области применения модели, разработку документированных требований и критериев оценки, а также формирование тестовых наборов данных, отражающих предполагаемую область применения.

Оценка производительности подразумевает демонстрацию адекватной точности и надёжности модели в реалистичных условиях с применением как количественных метрик (точность, полнота), так и качественного анализа результатов.

Устойчивость характеризует способность модели сохранять стабильную работу при вариациях и дрейфе данных (в т. ч. при случайных изменениях формата). Для её обеспечения необходимо тестирование на различных подвыборках (по источникам, регионам, языкам), а также проведение испытаний на устойчивость к намеренным искажениям входных данных.

Стабильность означает воспроизводимость результатов при повторных запусках на одних и тех же данных. Для её обеспечения необходимы повторные прогоны, ведение журналов параметров, верификация версии модели и оценка степени согласованности выходных данных для стохастических методов.

Адаптивность (возможность дообучения) — возможность обновления модели на основе новых данных. Данный аспект предполагает непрерывный мониторинг производительности, выявление дрейфа данных, наличие механизма дообучения на актуальных данных, а также разработку плана управления изменениями с переоценкой модели на основе эталонных наборов данных.

Неотъемлемым условием внедрения ИИ-моделей в процессы ФН является обеспечение их прозрачности [10]. Требуется открытое информирование о том, когда и каким образом ИИ-модели применяются для решения ключевых задач. При этом описание применяемых решений должно быть исчерпывающим: необходимо раскрывать архитектуру модели, характер ожидаемых входных и выходных данных, а также уровень и форму взаимодействия человека с ИИ-моделью.

Особой формой прозрачности является объяснимость ИИ-модели — возможность понять, каким образом система пришла к конкретному результату. Практическое значение объяснимости состоит в том, что она позволяет понять логику формирования ответа ИИ-модели и, при необходимости, оспорить полученный результат.

Вместе с тем необходимо учитывать существенное ограничение генеративных LLM: несмотря на то что такие модели могут формировать внешне убедительные обоснования своих результатов, эти обоснования представляют собой лингвистические конструкции, а не отражение реальных внутренних вычислений модели, и по своей сути являются постфактум рационализацией. В связи с этим для применения в ФН предпочтительно использование архитектуры Retrieval-Augmented Generation (генерация ответа, дополненная результатами поиска), которая не позволяет модели «фантазировать», а заставляет искать ответы строго в рамках изолированной базы документов. Ещё одним техническим решением для борьбы с галлюцинациями, ложной аргументацией и недетерминированностью LLM является привязка к источнику, т.е. требование предоставлять обоснования, привязанные к конкретным доказательствам (с ссылками на фрагменты входных данных или исходные документы). До включения сгенерированных LLM обоснований в аудиторский след рекомендуется проведение тестов на их достоверность [10].

Если говорить о качестве данных, то одним из требований к формированию обучающих данных можно считать структурную полноту каждого обучающего материала. Такой материал должен содержать заранее определённый набор обязательных элементов, и, при отсутствии хотя бы одного из них, обучающий материал признаётся непригодным для

включения в обучающую выборку независимо от качества остального содержимого.

Для минимизации эффекта переноса предвзятостей и ошибок человека возможно привлечение к оценке качества обучающих данных нескольких специалистов, а также специалистов с различным профессиональным профилем (например, клинических фармакологов, регуляторных экспертов и специалистов по безопасности лекарственных средств). Расхождение в оценках будет являться основанием корректировки обучающего материала перед его включением в обучающую выборку.

А вот для решения проблемы недостаточности количества данных для обучения ИИ-модели на текущий момент нет готового решения. Применение синтетических данных в ФН нельзя считать безопасным и надёжным по умолчанию. В связи с этим в ситуациях, когда объём доступных обучающих данных недостаточен, более обоснованным решением будет временный отказ от применения ИИ-модели в данном направлении с одновременным целенаправленным накоплением необходимых данных для последующего переобучения модели.

Обеспечение конфиденциальности данных при работе с ИИ-моделями предполагает реализацию комплекса мер. Согласно рекомендациям CIOMS [10], основными мерами защиты являются анонимизация и псевдонимизация данных. Практическим примером служит обучение алгоритмов на де-идентифицированных ИСНР, в которых персональные данные (инициалы и адреса пациентов и т.п.) заменены на условные идентификаторы или указан только возраст пациента в полных годах вместо точной даты рождения. Также при подготовке обучающих данных необходимо убедиться, что в них содержится только необходимая информация и отсутствует лишняя и ненужная для обучения модели.

К другим мерам защиты относятся использование закрытых локальных моделей, функционирующих в защищённом контуре, организация мониторинга запросов и данных, проведение регулярных аудитов вводимых и выводимых данных, применение политики хранения и удаления данных, закрепление требований к защите данных в договорных соглашениях с подрядчиками, проведения оценки воздействия на защиту данных, назначения ответственного сотрудника по защите данных, а также обеспечение регулярного обучения персонала и проведение аудитов [10].

Для успешной реализации всех перечисленных аспектов внедрения ИИ-модели недостаточно одного или двух человек. Требуется формирование многопрофильной команды. Успешная интеграция ИИ в процессы ФН требует привлечения междисциплинарной экспертизы на всех этапах жизненного цикла системы — от разработки до повседневной

эксплуатации. Это достигается посредством тесного взаимодействия специалистов различных направлений: ФН, обеспечения качества, науки о данных, информационных технологий, анализа платформ, этики, юриспруденции, защиты персональных данных и управления проектами [10].

Принципиально важно, чтобы сотрудники, непосредственно работающие, например, с ИСНР, принимали участие на этапах проектирования, тестирования и внедрения ИИ-моделей. Это является необходимым условием для обеспечения соответствия разрабатываемых решений реальным задачам и ожиданиям конечных пользователей.

Даже хорошо обученная и прошедшая валидацию ИИ-модель, разработанная на большом массиве вариативных данных, требует постоянного человеческого контроля. Уровень контроля над ИИ-моделью должен соответствовать потенциальным последствиям её ошибок или искажённых выводов [10]. В зависимости от характера наблюдаемой активности и степени автономии системы выделяют три формы человеческого надзора:

- *HITL (Human-in-the-Loop, человек в цикле)* — человек участвует в принятии решений на каждом этапе работы системы.
- *HOTL (Human-on-the-Loop, человек над циклом)* — человек вовлечён в проектирование и мониторинг системы, однако не участвует в каждом отдельном решении.
- *HIS (Human-in-Command, человек управляет)* — человек определяет стратегию применения ИИ, включая условия его использования и право отказать от него.

На начальных этапах внедрения предусматривается высокий уровень участия человека (HITL).

По мере накопления уверенности в стабильности работы системы степень вовлечённости специалиста по ФН в контроль качества выдаваемых решений может быть постепенно снижена.

Приведённые здесь аспекты по внедрению не являются исчерпывающими (особенно в части валидации), но могут служить основой для начального этапа внедрения ИИ-моделей в рутинные процессы ФН.

Закключение / Conclusion

ИИ-модели и, в частности, NLP-модели обладают значительным потенциалом для внедрения в рутинные процессы ФН, так как позволяют существенно сократить временные затраты на выполнение рутинных задач, повысить стандартизацию и качество документов, снизить влияние человеческого фактора. Всё это позволяет специалисту по ФН уделять больше времени экспертной аналитической работе.

Вместе с тем успешная интеграция ИИ-моделей невозможна без системного подхода к управлению сопутствующими рисками. Принципиально важно, что ИИ не замещает профессиональное суждение специалиста, а выступает инструментом его поддержки. Ответственность за принятые решения остаётся за человеком.

Внедрение ИИ-моделей в рутинные процессы ФН не является единовременным решением. Это непрерывный процесс, требующий междисциплинарного взаимодействия, чёткой спецификации задач, систематической валидации и поэтапного роста доверия к системе на основе верифицированных результатов. При соблюдении этих условий ИИ способен стать полноценным инструментом повышения эффективности процессов ФН.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов в связи с данной публикацией.

Участие авторов

Все авторы внесли значимый вклад в проведение исследования и подготовку статьи, прочли и одобрили финальную версию статьи перед публикацией, выразили согласие нести ответственность за все аспекты работы, подразумевающую надлежащее изучение и решение вопросов, связанных с точностью или добросовестностью любой части работы. Широбоков Я. Е. — поиск и анализ литературы, написание черновика рукописи; Молчанова Ю. А. — поиск литературы, написание статьи; Чумак Е. П. — редактирование статьи; Скрипкин А. Ю. — редактирование статьи.

ADDITIONAL INFORMATION

Conflict of interests

The authors declare that there is no conflict of interest in connection with this publication.

Authors' participation

All authors made a significant contribution to the research and preparation of the manuscript, read and approved the final version of the article before publication, and agreed to be responsible for all aspects of the work, including proper study and resolution of issues related to the accuracy or integrity of any part of the work. Shirobokov Y. E. — literature search and analysis, manuscript writing; Molchanova Yu. A. — literature search and analysis, manuscript writing; Chumak E. P. — manuscript editing; Skripkin A. Yu. — manuscript editing.

Финансирование

Исследование выполнено без целевого внешнего финансирования.

Financing

The study was carried out without targeted external funding.

СВЕДЕНИЯ ОБ АВТОРАХ

Широбокров Ярослав Евгеньевич — к. фарм. н., специалист группы экспертизы безопасности лекарственных средств департамента безопасности лекарственных средств АО «Р-Фарм», Российская Федерация

Автор, ответственный за переписку

e-mail: ye.shirobokov@rpharm.ru

ORCID ID: 0000-0002-7313-3432

РИНЦ SPIN-код: 5767-3000

Молчанова Юлия Андреевна — специалист группы экспертизы безопасности лекарственных средств департамента безопасности лекарственных средств АО «Р-Фарм», Российская Федерация

e-mail: ya.molchanova@rpharm.ru

ORCID ID: 0009-0009-5649-2159

Чумак Евгения Павловна — к. м. н., руководитель группы экспертизы безопасности лекарственных средств департамента безопасности лекарственных средств АО «Р-Фарм», Российская Федерация

e-mail: chumak@rpharm.ru

Скрипкин Алексей Юрьевич — директор департамента безопасности лекарственных средств АО «Р-Фарм», Российская Федерация

e-mail: skripkin@rpharm.ru

ABOUT THE AUTHORS

Yaroslav E. Shirobokov — Cand. Sci. (Pharm.), Specialist of Drug Safety Expertise Group, Drug Safety and Pharmacovigilance department, JSC R-Pharm, Russian Federation

Corresponding author

e-mail: ye.shirobokov@rpharm.ru

ORCID ID: 0000-0002-7313-3432

RSCI SPIN-code: 5767-3000

Yuliya A. Molchanova — Specialist of Drug Safety Expertise Group, Drug Safety and Pharmacovigilance department, JSC R-Pharm, Russian Federation

e-mail: ya.molchanova@rpharm.ru

ORCID ID: 0009-0009-5649-2159

Eugenia P. Chumak — Cand. Sci. (Med.), Head of Drug Safety Expertise Group, Drug Safety and Pharmacovigilance department, JSC R-Pharm, Russian Federation

e-mail: chumak@rpharm.ru

Aleksei Yu. Skripkin — Director of Drug Safety and Pharmacovigilance department, JSC R-Pharm, Russian Federation

e-mail: skripkin@rpharm.ru

Литература / References

1. Фоменко Н. М., Кузьмина Д. А., Миргород Е. П. Применение искусственного интеллекта в сфере государственного управления: риски, возможности, перспективы. *Вестник евразийской науки*. 2025;17(4):73FAVN425. [Fomenko N.M., Kuzmina D.A., Mirgorod E.P. Application of artificial intelligence in the field of public administration: risks, opportunities, prospects. *The Eurasian Scientific Journal*. 2025;17(4):73FAVN425 (In Russ.)].
2. Haykin S. *Neural Networks and Learning Machines*. 3rd ed. Pearson Education, 2009.
3. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press, 2016.
4. ГОСТ Р 71476-2024. Искусственный интеллект. Концепции и терминология искусственного интеллекта. М.: Российский институт стандартизации, 2024. 62 с. [GOST R 71476-2024. Artificial Intelligence. Concepts and Terminology of Artificial Intelligence. Moscow: Russian Institute of Standardization, 2024. 62 p. (In Russ.)].
5. Рахимов Т. Ю., Харисов А. Р. Исследование применимости нейросетевой архитектуры трансформера в детекции сетевых вторжений (DDoS, DrDoS). *Вестник науки*. 2026;4(97): 880-915. [Rakhimov T. Yu., Kharisov A. R. Study of the applicability of the transformer neural network architecture in detecting network intrusions (DDoS, DrDoS). *Science Bulletin*. 2026;4(97): 880-915. (In Russ.)].
6. Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. *Advances in Neural Information Processing Systems*. 2017. arXiv:1706.03762.
7. Габидуллина Л. Ф., Котловский М. Ю. Этика применения LLM-моделей в медицине и науке. *Медицинская этика*. 2025;(3):9–13. DOI: 10.24075/medet.2025.016. [Gabidullina L.F., Kotlovskiy M.Yu. Ethics of applying LLM-models in

- medicine and science. *Medicla Ethics*. 2025;(3): 9–13. doi: 10.24075/medet.2025.016 (In Russ.).
8. Goldberg Y. Neural Network Methods for Natural Language Processing. Morgan & Claypool Publishers, 2017.
9. Dreisbach C, Koleck TA, Bourne PE, Bakken S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *Int J Med Inform*. 2019 May;125:37-46. doi: 10.1016/j.ijmedinf.2019.02.008.
10. Artificial intelligence in pharmacovigilance. CIOMS Working Group XIV report. Geneva: Council for International Organizations of Medical Sciences (CIOMS), 2025.
11. Решение Совета ЕЭК № 87 от 3 ноября 2016 г. Об утверждении Правил надлежащей практики фармаконадзора Евразийского экономического союза. [Decision of the EEC Council No. 87 of November 3, 2016 On approval of the Rules of Good Pharmacovigilance Practice of the Eurasian Economic Union.]
12. Переверзев А.П., Миронов А.Н., Бунятян Н.Д., и др. Качественные методы определения отношения польза / риск фармакотерапии. *Безопасность и риска фармакотерапии*. 2014;(2):22-27 [Pereverzev A.P., Mironov A.N., Bunyatyan N.D., et al. Qualitative methods of benefit / risk assessment. *Safety and Risk of Pharmacotherapy*. 2014;(2):22-27. (In Russ.)].