

ИНФОРМАЦИОННЫЕ  
ТЕХНОЛОГИИ  
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION  
TECHNOLOGIES  
AND TELECOMMUNICATIONS

УДК 303.732.4

DOI: 10.17212/2782-2001-2025-1-27-48

## Проблемы разработки систем автоматизированной проверки договоров\*

**А.А. КОТОВ**

*РФ, 199178, г. Санкт-Петербург, ул. 14-я линия, 39, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский федеральный исследовательский центр Российской академии наук»*

*[alexanderkotovspb@gmail.com](mailto:alexanderkotovspb@gmail.com).*

Настоящая работа посвящена разработке системы автоматизированной проверки договоров и проблемам, возникшим в процессе ее реализации. Актуальность исследования заключается в решении задач оптимизации работы экспертов по проверке договоров.

В рамках исследования выделены ключевые проблемы в данной области: 1) отсутствие стандартизированной системы оценки рисков, что затрудняет принятие решений для лиц, принимающих решения; 2) недостаток открытых данных для обучения системы, ограничивающих ее способность эффективно выявлять ошибки и правильно интерпретировать юридические термины; 3) недостаточный контроль за процессом при использовании крупных языковых моделей, которые, несмотря на мощные аналитические возможности, не всегда могут гарантировать точность и согласованность выводов; 4) отсутствие учета контекста и семантической вариативности в юридических текстах затрудняет точную классификацию и выявление ошибок, особенно в случаях, когда одинаковые юридические термины или фразы могут иметь разные значения в зависимости от контекста.

По результатам исследования были получены: 1) авторская система классификации рисков юридических ошибок, основанная на теории нечетких множеств; 2) алгоритмы функционирования системы проверки юридических документов; 3) методы генерации данных для дообучения моделей BERT и корректировки моделей GPT.

Кроме того, в статье предложено использование комбинированного подхода с применением моделей GPT и BERT (полимодельная система) для оптимизации процесса: модель GPT обрабатывает текст и находит ошибки на основе заранее заданных подсказок, а модель BERT используется для присвоения ошибкам уровня риска и контроля над результатами работы модели GPT и обеспечивает более надежный и последовательный анализ.

Также акцентируется внимание на необходимости создания открытых датасетов с договорами, основанными на российском праве. Создание таких ресурсов не только улучшит качество обучения систем искусственного интеллекта, но и повысит доступность и прозрачность информации для исследователей и практиков в области права. Это станет важным шагом к совершен-

---

\* Статья получена 19 ноября 2024 г.

ствованию автоматизированного анализа юридических документов и увеличению общей правовой уверенности при их использовании.

**Ключевые слова:** автоматизированная проверка договоров, юриметрика, большие языковые модели, лигалтех, вычислительное право, анализ данных, искусственный интеллект, оценка рисков, стандартизация

## ВВЕДЕНИЕ

Актуальность разработки системы автоматизированной проверки юридических документов обусловлена необходимостью повышения эффективности и точности юридического анализа. В условиях роста объема правовой документации традиционные методы проверки становятся недостаточно быстрыми, что увеличивает вероятность ошибок, связанных с человеческим фактором. Автоматизация этого процесса с использованием искусственного интеллекта (ИИ) позволяет снизить количество ошибок, улучшить оценку рисков и ускорить обработку документов. Тем не менее, несмотря на очевидные преимущества, существует ряд серьезных вызовов, связанных с корректным распознаванием текста, подбором точных подсказок, а также обеспечением достоверности данных и унификации оценки рисков, что делает эту тему особенно актуальной для дальнейших исследований и разработки решений.

На рынке юридических технологий (legal-tech) существует широкий спектр решений, включая экспертные системы и конструкторы юридических документов, в которых пользователь заполняет необходимые поля или выбирает заранее подготовленные условия для договора. К более современным подходам относится автоматическая генерация и проверка документов с использованием больших языковых моделей (LLM) и методов обработки естественного языка, результаты которых впоследствии проверяет эксперт. Современные юристы отмечают, что использование технологий в их деятельности становится обязательным элементом успешной работы, а программы для управления договорами (contract management tools) являются ключевыми инструментами для этого [1].

В 1980-х годах юридические рассуждения (legal reasoning) стали одной из областей исследований в искусственном интеллекте. Тогда ученые полагали, что формальные модели юридических рассуждений могут способствовать развитию ИИ на основе правил, прецедентов и гипотетических рассуждений [2]. Однако со временем это направление утратило свою актуальность из-за сложности юридического анализа и ограниченных возможностей применения логического программирования [3, 4]. Кроме того, эмпирически было установлено, что основным ограничением развития экспертных систем является необходимость ручного наполнения базы знаний, что требует постоянного взаимодействия между экспертами и техническими специалистами и ограничивает возможности системы по самообучению и анализу больших данных [5].

На текущем этапе развития юриметрики – науки, изучающей применение вычислительных методов к правовым нормам [6], активно развивается направление «Вычислительное право» (computational law), ориентированное на автоматизацию юридического анализа (legal reasoning) [7]. Ключевой особенностью технологий, разрабатываемых в рамках этого направления, является

возможность применения нормативных актов к реальным или гипотетическим случаям с минимальным участием экспертов-юристов или без их участия. Дисциплина «Вычислительное право» уже преподается в таких крупных университетах, как MIT и Стэнфорд.

Развитие вычислительного права привело к созданию систем, использующих вероятностные суждения для решения юридических задач. Одним из ярких примеров таких систем являются программы автоматизированной проверки юридических документов (AI contract review software). Эти системы позволяют пользователям анализировать юридические документы на наличие рисков и нарушений, извлекать ключевую информацию и представлять ее в доступной форме, при этом не требуют от пользователей экспертных знаний в области права. Основой работы таких систем служат методы машинного обучения и обработки естественного языка.

## 1. ПОСТАНОВКА ЗАДАЧИ

Основная задача исследования – разработать прототип системы автоматизированной проверки договоров.

Пусть имеется система автоматизированной проверки договоров, которая принимает на вход юридический документ (D). Зададим следующие параметры и ограничения.

**Входные параметры** – юридический документ D, который должен быть обработан. Документ должен являться договором, составленным на русском языке, и регулироваться российским правом.

Форматы документа могут включать текстовые файлы (например, DOCX, TXT) или отсканированные изображения / сканированные PDF-файлы, для которых требуется применение оптического распознавания текста (OCR).

**Распознавание текста.** Необходимо выполнить распознавание текста в документе D, преобразовав его в текстовую форму T(D).

- В случае успешного распознавания документ передается для дальнейшей обработки.
- В случае если текст не может быть распознан, система должна вернуть ошибку ( $\emptyset$ ) – «Документ не может быть распознан».
- В случае если часть текста не распознана, система должна вернуть пользователю сообщение о том, что анализ договора проводится с учетом неточностей, возникших при распознавании текста.

**Поиск ошибок.** После успешного распознавания текста требуется провести анализ документа на наличие юридических ошибок E(T(D)). Если ошибки найдены, т. е.  $E(T(D)) \neq \emptyset$ , для каждой ошибки необходимо определить уровень риска R(E), который будет отражать потенциальную критичность ошибки в контексте юридической практики.

**Обработка ошибок.** Если ошибки найдены, необходимо предоставить информацию о них, включая их местоположение в документе и уровень риска, связанный с каждой ошибкой (с возможностью исключить определенный уровень риска из выдачи). В случае отсутствия ошибок в документе система должна сообщить об этом.

## 2. ОБЗОР ИССЛЕДОВАНИЙ ПО ТЕМАТИКЕ АВТОМАТИЗИРОВАННОЙ ОБРАБОТКИ ЮРИДИЧЕСКИХ ТЕКСТОВ

### 2.1. НАУЧНЫЕ ИССЛЕДОВАНИЯ ПО ТЕМАТИКЕ

В исследованиях, связанных с непосредственной разработкой систем проверки договоров, используются следующие методы в рамках программной реализации: 1) методы обработки естественного языка, 2) методы машинного обучения, 3) методы глубокого обучения. Ключевые актуальные исследования, посвященные непосредственной разработке систем для проверки положений договора, представлены в табл. 1. Они находятся в части предметной области, которая относится к строительной сфере, так как автоматизация строительных контрактов стала приоритетом из-за высокой сложности документации, множества нормативных требований, наличия большого количества стандартов, необходимости управления рисками и многосложных взаимодействий между сторонами, что требует повышения эффективности и минимизации ошибок.

Таблица 1

Table 1

#### Обзор систем обработки юридических документов

##### An overview of legal document processing systems

№ п/п	Тема исследования	Постановка проблемы	Решение	Результаты
1	Использование NLP для автоматизированной проверки контрактов и оценки рисков в строительных договорах [8]	Сложности с «глубокой» проверкой договора экспертом (короткий срок периода торгов)	Разработка автоматизированной модели анализа текста на основе естественного языка (NLP) и контролируемого машинного обучения (ML)	Максимальный результат f1 score (0,86) и accuracy (0,89) для классификации типа предложения и распределения риска для лица (0,76 и 0,83 соответственно)
2	Обнаружение рисков в строительных договорах при помощи больших языковых моделей [9]	Методы машинного обучения требуют качественных данных и больших усилий для разметки, чтобы эффективно использовать модель для проверки договоров. С другой стороны, языковые модели не обладают специфичными для сферы знаниями и могут «галлюцинировать», что требует определения способа их доннастройки	Построение системы проверки договоров на базе большой языковой модели с доннастройкой при помощи специальных знаний (строительного домена и промптов)	Результаты не поддаются адекватной количественной оценке, но эксперты заявляют повышение точности в проверке

Окончание табл. 1

End of the Tab. 1

№ п/п	Тема исследования	Постановка проблемы	Решение	Результаты
3	Автоматизированная проверка спецификаций к строительным договорам при помощи NLP [10]	Большое количество условий, отличающихся друг от друга в спецификациях (для разных стран), что затрудняет оперативный и единообразный анализ. Необходимо быстро выделять рискованные условия	Разработка автоматизированной системы для сравнения положений строительных спецификаций путем извлечения ключевых положений с помощью методов NLP в машинном обучении	Ключевые положения извлекаются из текста (F1 score – 0,928)
4	Применение RAG-модели для анализа договорных рисков в строительных проектах [11]	Существующие методы анализа контрактов часто требуют значительных усилий и времени. Авторы разработали автоматизированную систему, в которой используются методы обработки естественного языка (NLP) и извлечение данных, чтобы улучшить точность и эффективность анализа контрактов	Разработка системы С3RAG, которая сочетает методы обработки естественного языка (NLP) и извлечения данных для автоматизированного анализа рисков контрактов в строительных проектах	Читабельность (0,91), полученные данные из дел (0,74) и точность цитирования (0,87)
5	Автоматизированное разделение требований контракта на проектирование и строительство для поддержки составления субконтрактов [12]	Подрядчикам приходится разрабатывать несколько субподрядных договоров, включающих лишь часть требований основного контракта. Ручная практика составления субподрядных договоров связана с ошибками и отнимает много времени	Модель классификации с использованием машинного обучения для классификации требований по трем заранее определенным категориям, включая проектирование, строительство, эксплуатацию и техническое обслуживание	Максимальный результат precision 93,20 %, recall 93,08 %, F-score 92,98 %

1. Для обучения системы Gorkem Eken, Irem Dikmen, M. Talat Birgonul [8] выбрали стандартные формы договоров FIDIC, которые применяются в международной практике строительства. Все предложения были размечены по типам (Risk, Right, Obligation, Heading, Definition) и принадлежности к риску (подрядчик, заказчик, совместный риск), что позволило создать обучающую базу данных из 1217 уникальных предложений вместо 5346 благодаря проверке на уникальность. Предобработка данных включала лемматизацию и составление списка стоп-слов, а разметку текстов проводили эксперты с более чем 10-летним опытом. Созданный датасет использовался для обуче-

ния и оценки 12 моделей машинного обучения с применением методов векторного представления слов, таких как «мешок слов» и TF-IDF. Наиболее эффективными сочетаниями оказались 1) TF-IDF и Support Vector Machine, 2) GloVe Embedding и RNN, 3) Word Embedding и RNN, а также их комбинации.

2. Для создания системы на основе большой языковой модели, обогащенной знаниями из строительного домена, Saika Wonga, Chunmo Zheng, Xing Su, Yinqiu Tang [9] выполнили четыре шага: 1) создание базы знаний, 2) извлечение клауз из договора, 3) оценка каждой клаузы экспертом, 4) разработка подходящих промптов. В первом шаге тексты договоров были разделены на части и представлены в векторном виде для сохранения в базе знаний, где элементы связаны чек-поинтами, клаузами и оценками. Во втором шаге использовались чек-листы от экспертов для извлечения клауз, при этом учитывались сложности, связанные с формулировкой правил в различных местах договоров. В третьем шаге собирались оценки клауз с использованием гибридного подхода для извлечения связанных топ- $k$  клауз и объединения их в векторное представление. На последнем этапе была разработана двухуровневая система промптов, которая генерирует несколько ответов на заданные вопросы и позволяет модели выбрать наиболее подходящий.

3. Для создания системы, находящей высокорисковые положения в спецификациях к строительным договорам, коллектив авторов (Seonghyeon Moon, Gitaek Lee, Seokho Chi) разработал систему на NLP-методах для создания тезауруса, классификации ключевых слов и векторного представления частей документа. В системе реализованы NER-модель с Bi-LSTM и CRF, а также Provision-pairing-модель, которая использует Doc2Vec для создания векторных представлений фрагментов текста и находит наиболее похожие по смыслу положения. В исследовании были подготовлены 56 спецификаций, которые перевели из pdf-формата в txt-формат, в результате было подготовлено 19 338 предложений для анализа. При помощи n-gram была проведена токенизация, затем составлен список стоп-слов и выполнена лемматизация. Затем авторы оценили векторы слов при помощи косинусного сходства и убрали те, где показатель ниже 0,5. Кроме того, для решения проблемы рекурсивной замены авторы предложили использовать алгоритм PivotTermDetermination, через который определяется наиболее важное слово по максимальному количеству отсылок на него от других схожих слов и минимальному количеству отсылок этого слова на другие слова. Для того чтобы учитывать контекст ключевых слов в больших документах, где необходимые слова могут быть разъединены, использовалась комбинация методов Bi-LSTM и CRF. Также опросили 25 экспертов для определения слов, являющихся маркером рискованных положений. В качестве единицы для извлечения из текста авторы решили извлекать блок по конкретной тематике вместо наиболее подходящего документа (слишком широко) или подходящего предложения (слишком узко). Полученные результаты позволили эффективнее проводить проверку положений из разных договоров и минимизировать субъективность проверки для риск-менеджмента.

4. В исследовании Shuai, Bing and Caldas, Carlos H. [1] была предложена система C3RAG, основанная на методологии, сочетающей обработку естественного языка и извлечение данных для анализа рисков контрактов в строительных проектах. Модель была обучена на большом наборе данных из реальных

судебных решений с использованием методов извлечения и генерации информации, включая смешанные модели (RAG), чтобы обеспечить надежный анализ рисков, основанный на предшествующих случаях, с целью автоматизации и повышения точности процесса анализа контрактов. Результаты показывают, что система C3RAG эффективно обрабатывает и анализирует контракты, достигая высоких оценок по метрикам читабельности (0.91), получения данных из дел (0.74) и точности цитирования (0.87).

5. Fahad ul Hassan, Tuyen Le [12] разработали многоклассовую модель для классификации требований проекта в договорах, разделяя их на три категории: проектирование, строительство и эксплуатация. Для упрощения задачи были исключены положения, не относящиеся к требованиям, и принято допущение, что каждое требование относится только к одной категории. Для обучения использовались два датасета: размеченный и не размеченный экспертами. Тексты предобрабатывались с использованием приведения к нижнему регистру, токенизации, удаления стоп-слов и лемматизации. Векторизация осуществлялась с помощью методов Bag of Words и Word Embeddings (например, word2vec и skip-gram), что позволило обработать 362,284 предложения. Для обучения моделей были установлены частотный порог, размер окна и размер вектора. Средние векторы рассчитывались с использованием метода TF-IDF, а выбор признаков осуществлялся с помощью методов  $\chi^2$ , Mutual Information и RFE. Применялись пять методов машинного обучения (NB, SVM, LR, DT и kNN) с использованием 10-кратной кросс-валидации для оценки производительности. Гиперпараметры оптимизировались с помощью тестирования различных комбинаций параметров для каждого алгоритма, включая настройку регуляризации для SVM и LR, и выбор критерия Джини для DT. Для глубоких моделей (CNN и RNN) использовались последовательности слов с сохранением порядка слов в предложениях. Для повышения устойчивости и точности моделей применялись ансамблевые методы: случайный лес, бэггинг, бустинг и голосование. Эффективность классификации оценивалась метриками полноты, точности и F-меры.

Также существуют обзорные исследования по этой тематике. А. Канапала и коллеги [13] рассмотрели современные методы резюмирования юридических текстов, включая одиночные и множественные документы, протестированные на различных юридических наборах данных. Методы были разделены на четыре категории: лингвистические, графовые, семантические и классификационные. Артур Д. [14] предоставил обзор методов машинного обучения для анализа юридических текстов, включая переход от традиционных подходов, таких как «мешок слов», к более современным, как BERT, который используется для классификации документов и прогнозирования судебных разбирательств и показал высокую эффективность. Н. Бансал и другие [15] провели исследование применения глубокого обучения (CNN, RNN, LSTM, GRU) для различных юридических задач, включая классификацию и прогнозы, выявив, что эти модели показывают высокие результаты в анализе юридических текстов.

## 2.2. ОБЗОР ОСНОВНЫХ СИСТЕМ АВТОМАТИЗИРОВАННОЙ ПРОВЕРКИ ЮРИДИЧЕСКИХ ДОКУМЕНТОВ

Для начала осуществим краткий обзор систем автоматизированного анализа договоров на примере следующих программ: Blackboiler, Foley Equipped, Legly, GoHeather, Genie AI, Embedika, Doczilla, Noroots.

Системы автоматизированной проверки юридических документов используют сочетание обработки NLP и ML, а иногда и алгоритмов глубокого обучения для понимания контекста, выявления ключевых элементов и оценки рисков в договорах. Автоматизируя извлечение необходимой информации, такой как пункты, даты истечения срока действия договора и обязательства, инструменты анализа договоров на основе ИИ значительно сокращают время и усилия, необходимые специалистам по правовым вопросам для работы с договорами.

Функциональные возможности таких систем включают семантический поиск, который позволяет пользователям находить положения на основе смысла их запросов, а не на точном совпадении ключевых слов. Это особенно полезно в контексте юридических документов, где одно и то же понятие может быть сформулировано по-разному в разных документах. Кроме того, некоторые системы предлагают предиктивную аналитику, позволяющую прогнозировать потенциальные юридические проблемы до их возникновения, что обеспечивает проактивное управление рисками, вытекающими из договоров. Еще одной важной функцией является проверка соответствия, когда программное обеспечение сравнивает условия договора с соответствующими законами и правилами, чтобы убедиться в их соблюдении, и выделяет любые несоответствия для дальнейшего рассмотрения.

В этих системах могут использоваться методы обучения без учителя, но преимущественно используется обучение с учителем для классификации и извлечения информации, при этом модели обучаются на наборе данных договоров с метками, чтобы учиться на прошлых примерах. Современные системы могут использовать модели-трансформеры, такие как BERT (Bidirectional Encoder Representations from Transformers), для более глубокого понимания языка договоров с учетом их способности обрабатывать слова в связи со всеми остальными словами в предложении. Это позволяет более тонко интерпретировать юридические тексты, улавливая нюансы, которые традиционные системы, основанные на правилах, могут упустить.

Например, у одного из самых успешных коммерческих продуктов в этой области (Blackboiler Inc.) зарегистрирована система патентов, которая направлена на юридическую защиту технических решений, используемых для автоматизированного анализа юридических документов. Среди них: 1) метод предложения пересмотра анализируемого документа посредством распознавания фрагмента текста и его сравнения с информацией, хранящейся в базе данных; 2) системы и методы для обработки фраз в юридических документах посредством идентификации фрагмента текста и статистического подсчета эквивалентности положений; 3) системы и методы слотовой нормализации текстовых данных посредством распознавания изменяемого пункта документа, присвоения ему значения и автоматического обновления одного или нескольких пунктов; 4) система редактирования параметров посредством интерактивных элементов ввода [16–18].

Теперь перейдем к проблемам, с которыми столкнулся автор при разработке системы автоматизированной проверки договоров.

### **3. ПРОБЛЕМЫ И РЕШЕНИЯ**

#### **3.1. ЗАКРЫТОСТЬ ТЕКУЩИХ РЕШЕНИЙ**

Первая проблема разработки заключается в том, что код ключевых систем или его часть не находится в открытом доступе из-за коммерческой ценности. В результате этого исследователи не могут точно воспроизвести передовые системы, которые активно используются на практике, а также не могут проверить логику работы таких систем. Это ограничивает возможности для исследования, улучшения и адаптации уже готовых технологий. Минусы закрытых систем также включают отсутствие прозрачности и зависимость от поставщика в случае возникновения проблем с системой.

Для решения этой проблемы был разработан алгоритм функционирования системы, представленный на рис. 1, исходя из описаний функционала существующих коммерческих систем. Алгоритм включает ключевые этапы, такие как распознавание текста документа, поиск ошибок и классификацию рисков, что позволяет создать модель, максимально приближенную к передовым решениям. Такой подход не только устраняет зависимость от закрытых систем, но и предоставляет возможность гибко адаптировать алгоритм под специфические требования, такие как различные типы юридических документов или особенности законодательства.

Кроме того, использование открытых данных и доступных инструментов для построения системы позволяет исследователям и разработчикам проверять и совершенствовать алгоритм, а также проводить независимые тестирования, обеспечивая прозрачность процессов. В будущем эта система может быть доработана с учетом реальных потребностей пользователей и интегрирована с другими инструментами в области юридического анализа.

В рамках подготовки к реализации системы и ее непосредственной реализации на основе предложенного алгоритма автор также столкнулся со следующими проблемами.

#### **3.2. ОТСУТСТВИЕ ЕДИНООБРАЗНОЙ СИСТЕМЫ ОЦЕНКИ РИСКОВ**

Для эффективной автоматизации оценки рисков в договорах необходима предварительная стандартизация с разработкой четких семантических правил для выявления ошибок и оценки их потенциальных последствий. На данный момент отсутствуют общепризнанные стандарты ранжирования рисков, связанных с ошибками в договорах, что затрудняет создание качественных автоматизированных систем проверки договоров.

Под единообразной системой оценки рисков понимается набор правил и критериев, которые определяют, как именно идентифицировать и классифицировать ошибки, а также оценивать их влияние на юридические, финансовые и репутационные аспекты договора. Эти правила должны включать как формализованные методы (например, использование семантических моделей для анализа текста), так и стандартизированные шкалы для оценки уровня риска.

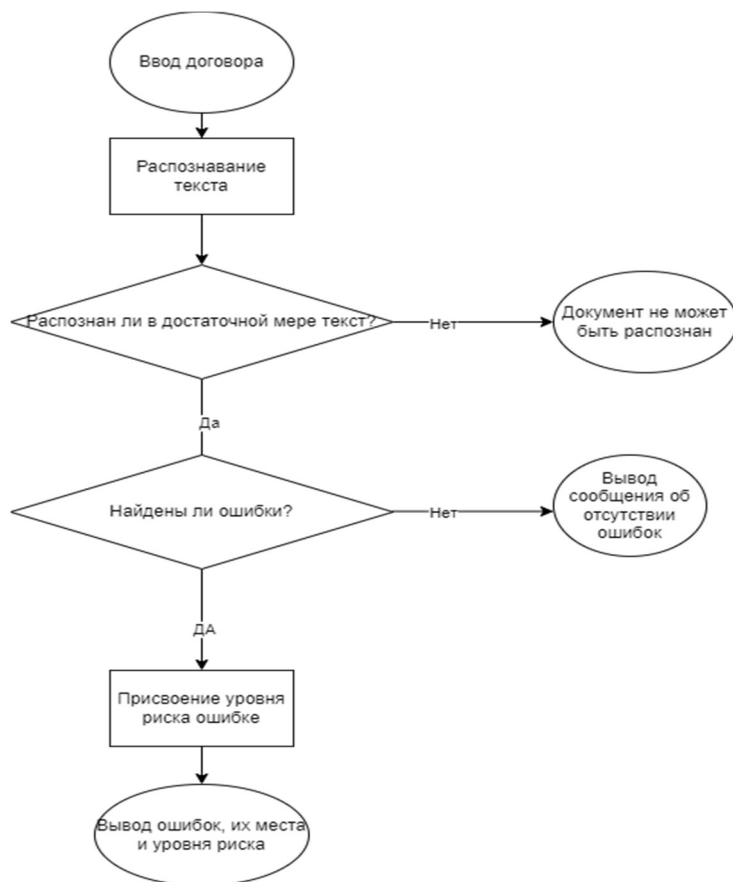


Рис. 1. Алгоритм проверки юридического документа при помощи системы автоматизированной проверки юридических документов

Fig. 1. An algorithm for verifying a legal document using an automated verification system for legal documents

Исследование Melvin A. Eisenberg [19], хотя и классифицирует ошибки на механические, интерпретационные и транскрипционные, не предоставляет необходимых инструментов для автоматизации, таких как алгоритмически определенные критерии оценки рисков. Это затрудняет разработку эффективной системы, поскольку эксперты опираются в основном на интуитивные суждения при определении уровня рисков, связанных с ошибками, что зачастую носит субъективный характер и изменяется в зависимости от конкретного контекста.

Необходимо провести углубленный анализ ошибок в договорах и их влияния на риски для разработки объективных критериев и семантических правил их оценки. Для этого необходимо исследование типичных ошибок, случаев, когда они приводили к юридическим последствиям, и привлечение экспертов для формирования базы знаний. Такие данные помогут улучшить автоматизированную систему и снизить вероятность субъективных ошибок, обеспечивая более точную оценку рисков. Ключевым этапом будет являться разработка системы, интегрирующей мнения экспертов и практические данные для создания алгоритмов, способных распознавать и классифицировать ошибки по установленным семантическим правилам. При этом необходимо учитывать

возможные расхождения мнений экспертов по вопросам наличия ошибки и уровня риска, который она вызывает.

В связи с этим автор решил выработать собственный инструментарий для оценки рисков. Учитывая нечеткость и небинарность знаний о том, какой риск может повлечь та или иная ошибка, адекватным решением задачи является адаптация подхода, предложенного А.В. Спесивцевым [20], в рамках которого каждому риску присваивается значение от низкого к высокому. В основе этого подхода лежит лингвистическая переменная Заде, которая позволяет оценивать риски с использованием нечетких терминов. Каждому риску присваивается степень принадлежности к одному из лингвистических терминов, описанных шкалой от низкого значения к высокому с помощью функции принадлежности:

$$\mu_{\text{Риск}}(x) \in [0, 1],$$

где  $\mu_{\text{Риск}}(x)$  – степень принадлежности ошибки  $x$  к категории риска (например, низкий, средний, высокий), а значения из интервала  $[0, 1]$  отражают степень этой принадлежности.

Получилось следующее формальное описание (табл. 2).

Таблица 2

Table 2

**Уровни риска и их описание**  
**Risk Levels and Their Description**

Уровень риска	Характеристики
Низкий риск $0 < \mu_{\text{Риск}} \leq 0.2$	Технические недочеты, которые не влияют на действительность договора, восприятие сторонами смысла условий договора, исполнение обязательств сторонами и реализацию прав
Ниже среднего $0,2 < \mu_{\text{Риск}} \leq 0.4$	Незначительные семантические (смысловые) ошибки, которые не влияют на действительность договора, восприятие сторонами смысла условий договора, исполнение обязательств сторонами и реализацию прав
Средний риск $0,4 < \mu_{\text{Риск}} \leq 0.6$	Ошибки, которые не влекут прямых негативных последствий, но могут ввести стороны в заблуждение относительно правовых последствий, порожаемых условиями договора (например, стороны могут прийти к разному представлению об условиях в силу их отсутствия в договоре)
Выше среднего $0,6 < \mu_{\text{Риск}} \leq 0.8$	Ошибки, которые могут привести к неверному восприятию сторонами условий договора и которые могут повлиять на исполнение сторонами своих обязательств, реализацию прав (исполнение обязательства не тому лицу, исполнение не того обязательства)

Окончание табл. 2

End of the Tab. 2

Уровень риска	Характеристики
Высокий риск $0,8 < \mu_ \leq 1.0$	Ошибки, которые могут повлечь признание договора незаключенным, взыскание компенсации за нарушение исключительных прав, взыскание штрафа за нарушение договора (неустойка), судебные разбирательства

Для апробации шкалы был проведен анализ 50 лицензионных договоров на предоставление программного обеспечения, в ходе которого выявлено 46 ошибок. Эти ошибки были классифицированы на четыре категории: 1) ошибки-пропуски, 2) ошибки-опечатки (изначальные ошибки, присутствующие в текстовом представлении документа, такие как неправильно написанные слова, грамматические или синтаксические недочеты. Они не связаны с этапом распознавания текста (OCR)), 3) ошибки – противоречия закону, 4) ошибки – существенные нарушения интересов одной из сторон. На основе выявленных ошибок был проведен опрос, в котором приняли участие 11 экспертов-юристов, использующих предложенную шкалу оценки рисков. Результаты опроса можно найти в [21]. В целом, эксперты продемонстрировали согласованность в оценках с преобладанием доминирующего мнения по большинству вопросов, что показано на рис. 2 и 3.

19. Неверно определен предмет договора  
11 ответов

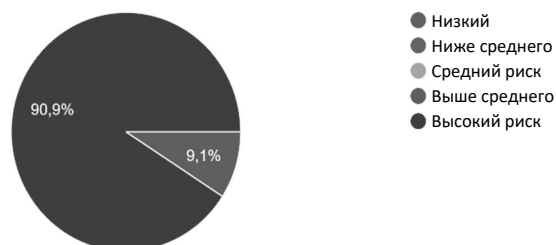


Рис. 2. Результаты ответов на вопрос о квалификации риска

Fig. 2. The results of the responses to the question on risk qualification

На диаграмме, представленной на рис. 4, отражено распределение голосов по присвоению уровня риска для ошибки, связанной с отказом от личных нематериальных авторских прав. Эксперты распределили свои ответы следующим образом:

- низкий риск: 18,2 % респондентов;
- ниже среднего риска: 18,2 % респондентов;
- средний риск: 9,1 % респондентов;
- выше среднего риска: 27,3 % респондентов;
- высокий риск: 27,3 % респондентов.

13. Не указан размер лицензионного вознаграждения  
11 ответов

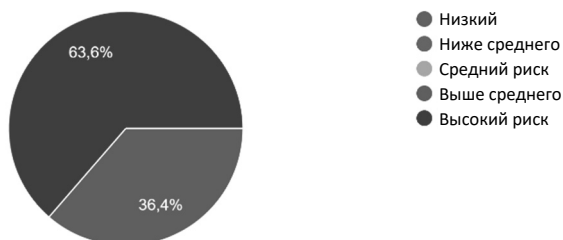


Рис. 3. Результаты ответов на вопрос о квалификации риска

Fig. 3. The results of the responses to the question on risk qualification

Но имеются также ситуации, когда у экспертов нет единого доминирующего мнения (рис. 4).

20. Содержится условие об отказе от личных неимущественных авторских прав  
11 ответов

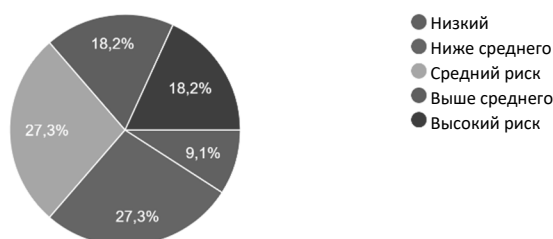


Рис. 4. Результаты ответов на вопрос о квалификации риска

Fig. 4. The results of the responses to the question about risk qualification

Из диаграммы видно, что голоса распределены достаточно равномерно, с наибольшим числом голосов, сосредоточенных на уровнях «Выше среднего» и «Высокий риск», что указывает на разногласия среди экспертов и отсутствие доминирующего мнения.

Результаты этого опроса будут подробно рассмотрены с точки зрения логики нечетких множеств, применяемой для анализа экспертного знания. В рамках настоящего исследования достаточно подтверждения того, что такая шкала может быть использована экспертами в целом.

### 3.3. НЕХВАТКА ДАННЫХ ДЛЯ ОБУЧЕНИЯ СИСТЕМЫ

Третья проблема, с которой столкнулся автор при разработке, заключается в отсутствии датасетов из юридического домена для обучения в открытом доступе. Эта проблема характерна для российского правопорядка, так как зарубежные авторы провели достаточную публикацию датасетов, которые могут использоваться для обучения любым из разработчиков. Например, Contract Understanding Atticus Dataset (CUAD): создан в 2020 году The Atticus Project, включает более 500 договоров (25 типов) с 13 101 аннотацией для обучения

систем автоматизированного анализа договоров [22]; Cornell eRulemaking Corpus (CDCP): состоит из 731 комментария пользователей и 4931 аннотации элементарных единиц, используется для систем извлечения аргументов из текста [23]; ACL/COLING 2014 Dataset: Корпус из 1010 политик конфиденциальности с популярных сайтов, собранных в 2013–2014 годах [24]; Privacy Q&A Corpus (EMNLP 2019): включает 1750 вопросов по политикам конфиденциальности с экспертными аннотациями [25]; ContractNLI: содержит 607 аннотированных контрактов для задач на основе естественного языка [26].

Отсутствие открытых датасетов в юридическом домене для обучения систем автоматизированной проверки договоров представляет собой серьезную преграду для разработчиков в России. Из-за специфики российского правового порядка и различий в правовых системах даже переведенные зарубежные данные не могут быть эффективно адаптированы для решения локальных задач. Это приводит к необходимости создания собственных датасетов, что требует значительных ресурсов и времени. Разработчикам необходимо сотрудничать с юридическими экспертами для сбора, аннотирования и классификации данных, чтобы обеспечить высокое качество и актуальность полученной информации.

Одним из возможных решений может быть инициирование проектов по совместному сбору данных, например, через коллаборацию с юридическими фирмами или образовательными учреждениями, что позволит создать репрезентативные датасеты, соответствующие российским реалиям. Также стоит рассмотреть возможность использования методов активного обучения, в которых система может запрашивать у экспертов разметку для неопределенных или сложных случаев, что снизит нагрузку на специалистов и ускорит процесс аннотирования.

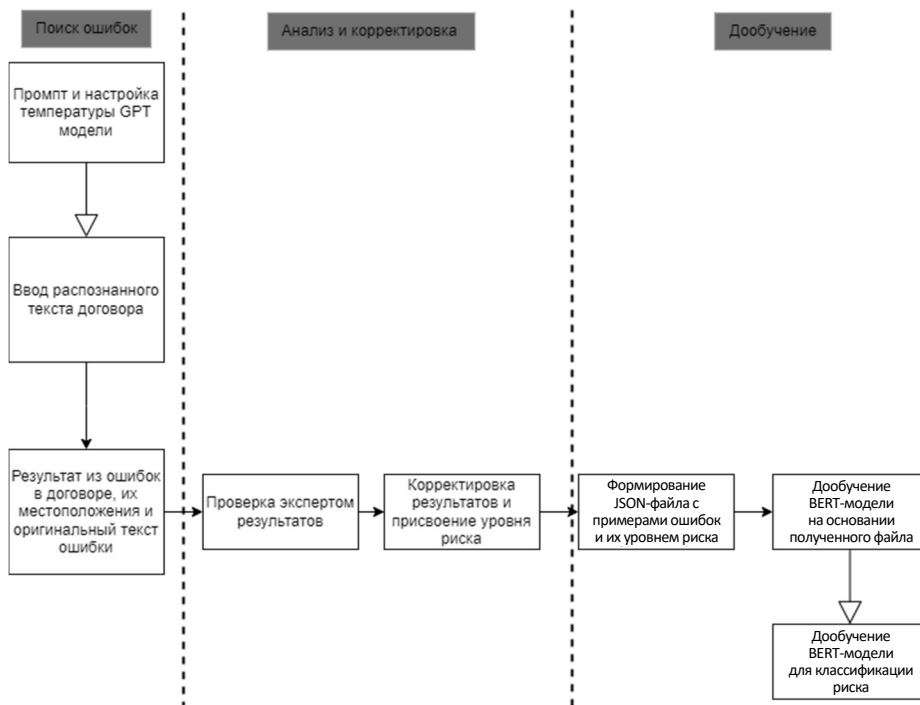


Рис. 5. Схема метода генерации данных для обучения

Fig. 5. The scheme of the data generation method for training

В рамках создания прототипа системы автор пришел к следующему решению. В силу отсутствия большого количества размеченных данных, которые автор смог бы использовать для обучения системы (обучение с учителем), наиболее рациональным инструментом для работы с текстом являются GPT-модели (например, gpt-3 turbo, gpt-4o-mini, yandexgpt-lite, LLaMMA). Результат работы этих моделей можно использовать для генерации данных при обучении и формировании датасетов с учетом предварительной обработки экспертом. Схема предлагаемого метода представлена на рис. 5.

Предложенный метод включает несколько этапов автоматизированного анализа юридических документов с использованием моделей искусственного интеллекта. На первом этапе GPT-модель принимает документы в формате Word или PDF, анализирует их в соответствии с заданным промптом на основе примеров ошибок из опроса и описания ошибок из табл. 1, а также температурой модели для поиска ошибок в тексте. На втором этапе эксперт проверяет результаты, корректирует их и присваивает каждой ошибке уровень риска, после чего сохраняет данные в JSON-формате. На последнем этапе JSON-файл используется для дообучения модели BERT, которая затем применяется для дальнейшей классификации ошибок по уровням риска и для улучшения точности системы. Этот подход позволяет интегрировать экспертные знания в процесс обучения и повысить эффективность автоматического анализа юридических документов.

#### **3.4. НЕДОСТАТОЧНЫЙ КОНТРОЛЬ ЗА ПРОЦЕССОМ ПРИ ИСПОЛЬЗОВАНИИ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ**

Автор также столкнулся с проблемой, при которой система при каждом запуске анализа одного и того же документа меняла формулировки, несмотря на то что суть их оставалась примерно одинаковой. Модель каждый раз присваивала разную степень риска, что делало ее выводы крайне ненадежными. Это подчеркивает одну из главных проблем, связанных с использованием больших языковых моделей, – их склонность к непредсказуемым и неконтролируемым результатам, что особенно важно в юридическом контексте.

В отличие от традиционных систем, основанных на правилах, современные модели машинного обучения, особенно глубокие нейросети, работают как «черные ящики» [27]. Это означает, что процессы принятия решений не прозрачны, и это может привести к трудностям в объяснении или обосновании выводов модели. Неправильно истолкованное положение или упущенная из виду оговорка, обусловленные предвзятостью или ограниченностью обучающих данных, могут иметь серьезные юридические последствия, такие как риск наложения штрафов, расторжение договора или взыскание денежных компенсаций. Кроме того, большие языковые модели могут «галлюцинировать» [28] и изображать следование промпту (Alignment faking) [29].

Использование модели GPT для автоматизированной проверки документов имеет свои ограничения. Одним из них является то, что модель не позволяет строго контролировать правила, на которых основан процесс анализа. Это может привести к неоднозначностям в выявлении ошибок и неэффективности работы системы, поскольку модель работает по принципу «черного ящика» и не всегда может объяснить свои выводы. Поскольку модель GPT, как правило, генерирует текст на основе вероятностных вычислений, это не

гарантирует стабильно точного результата, особенно если запросы подаются в разных формах. В связи с этим автор считает, что для повышения точности и надежности проверки документов может быть полезно комбинировать использование GPT-модели с BERT-моделью, которая будет выполнять наиболее важную часть – присваивать уровень риска для каждой из ошибок (рис. 6).

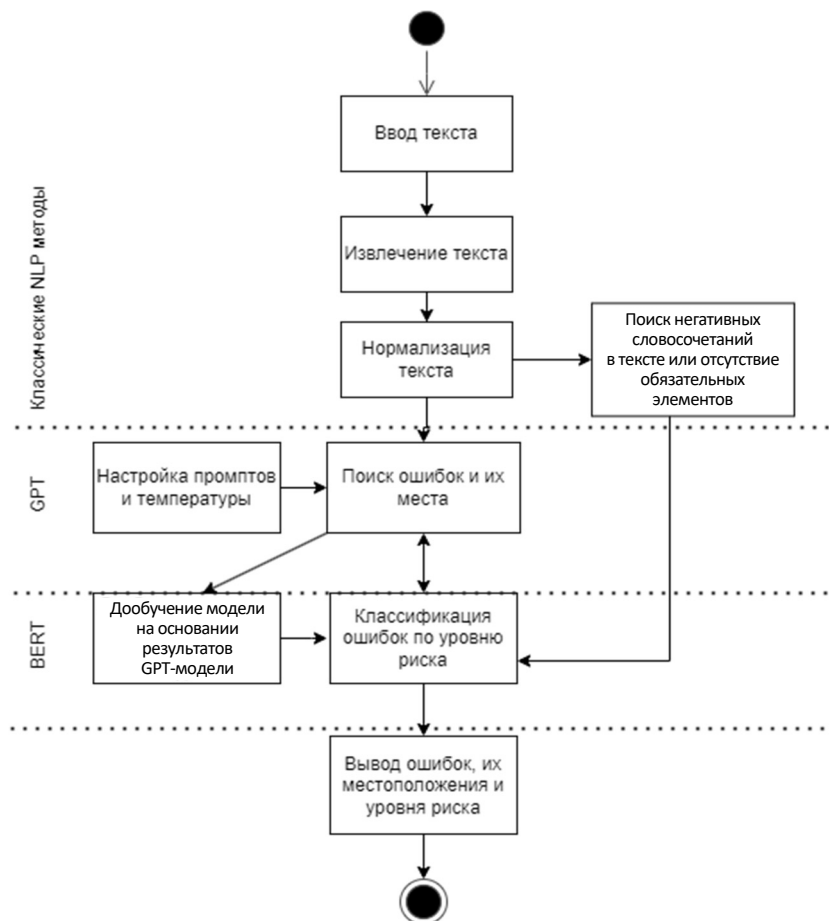


Рис. 6. Полимодельный комплекс для проверки договоров

Fig. 6. A Poly-Model System for Contract Verification

Для более подробного описания предложенной схемы (рис. 6) полимодельного комплекса проверки юридических документов можно выделить следующие ключевые этапы.

1. **Ввод текста.** На этом шаге пользователь вводит документ, который должен быть проанализирован. Это может быть текст в формате PDF, Word или в других форматах. Важно, чтобы текст был корректно извлечен для дальнейшей обработки.

2. **Извлечение текста.** Выполняется извлечение фактического текста из документа, если это не было сделано ранее. Если документ в формате PDF или в виде изображения, то может быть использовано распознавание текста (OCR), чтобы получить текстовый вариант документа для дальнейшего анализа.

3. **Нормализация текста.** Приведение текста к унифицированной форме для дальнейшего анализа: удаление ненужных символов, исправление пробелов, преобразование всех символов в нижний регистр и другие этапы предобработки, которые обеспечат корректную работу модели.

4. **Поиск ошибок с использованием GPT-модели.** Текст, подготовленный на предыдущем шаге, передается в GPT-модель. Важным параметром для этой модели является **температура** – параметр, который контролирует степень случайности или креативности в ответах модели. Чем выше температура, тем более разнообразными и менее предсказуемыми будут выводы модели. Температура влияет на то, насколько разнообразными и новыми могут быть ответы, что полезно для поиска ошибок в тексте, которые могут быть представлены различными способами [30]. Модель использует заданный **промпт**, то есть инструкции или вопросы, чтобы найти ошибки в тексте на основе примеров, предоставленных в предыдущем обучении или настройке модели.

5. **Проверка и классификация ошибок с использованием BERT-модели.** Результаты, полученные от GPT-модели, передаются в BERT-модель, которая используется для классификации ошибок, для более точной оценки и проверки найденных ошибок. Она помогает определить, насколько каждая ошибка важна, и классифицирует ее по уровням риска.

6. **Поиск негативных словосочетаний и отсутствующих элементов.** Для улучшения качества анализа дополнительно используется инструмент поиска негативных словосочетаний или пропущенных элементов, которые могут быть важны для содержания договора. Эта задача выполняется с помощью регулярных выражений и нечеткого поиска.

7. **Вывод ошибок, их местоположения и уровня риска.** На заключительном этапе система выдает отчет, который включает в себя список всех найденных ошибок, их местоположение в тексте и степень риска, присвоенную каждой ошибке. Это позволяет юристам и экспертам легко ориентироваться в документе и понимать, какие именно ошибки являются наиболее критичными.

Результаты работы системы и проблемы с ее оценкой станут предметом следующих исследований, включая необходимость применения ограничений и коэффициентов для уточнения шкалы рисков. Продемонстрируем часть результатов проверки типичного договора с указанием места ошибки, описания ошибки и уровня риска, который она порождает (табл. 3).

Таблица 3

Table 3

### Результаты работы системы автоматизированного анализа договора

#### Results of the Automated Contract Analysis System

Ошибка	Текст положения	Описание	Уровень риска
Отсутствуют полные реквизиты сторон	«ООО “_____”», «ИП _____»	Не указаны ИНН, ОГРН, юридические и банковские реквизиты	Высокий

Окончание табл. 3

End of the Tab. 3

Ошибка	Текст положения	Описание	Уровень риска
Несоответствие дат в договоре и приложении	Дата заключения договора: «25 октября 2024 г.»  Дата заключения договора в Приложении № 1: «от 19 февраля 2024 г.»	Несоответствие дат может вызвать споры о начале действия договора	Средний
Несоразмерная пеня за просрочку оплаты	<i>«Поставщик вправе требовать от него уплаты пени в размере 5 % от суммы задолженности за каждый день просрочки»</i>	Согласно ст. 333 ГК РФ суд может снизить неустойку, если она явно несоразмерна последствиям нарушения	Высокий
Применение Международного коммерческого арбитражного суда при ТПП РФ	<i>«Все споры подлежат разрешению в Международном коммерческом арбитражном суде при ТПП РФ»</i>	Если обе стороны – российские компании, то целесообразнее выбрать арбитражный суд по месту регистрации поставщика или покупателя	Средний

## ЗАКЛЮЧЕНИЕ

В настоящей статье рассмотрены ключевые проблемы, с которыми автор столкнулся при разработке системы автоматизированной проверки юридических документов. Основными проблемами являются отсутствие единой системы оценки рисков, нехватка данных для обучения и недостаточный контроль за процессом при использовании больших языковых моделей.

Для решения этих проблем была предложена авторская система классификации рисков, основанная на логике нечетких множеств. Этот подход позволяет учитывать неопределенности, присущие юридическим текстам, и более точно оценивать риски, связанные с возможными ошибками. Интеграция модели GPT с моделью BERT может обеспечить более единообразную оценку рисков и повысить надежность выводов системы.

Кроме того, исследование подчеркивает необходимость разработки открытых датасетов, посвященных договорам по российскому праву. Создание таких ресурсов не только позволит улучшить процесс обучения моделей искусственного интеллекта, но также повысить доступность и прозрачность информации для специалистов в области права.

## СПИСОК ЛИТЕРАТУРЫ

1. 2023 Legal Technology Report: for In-house Legal Professionals. – Association of Corporate Counsel, 2023. – URL: [https://www.acc.com/sites/default/files/2023-05/ACC\\_LegalTech-Report23\\_Final.pdf](https://www.acc.com/sites/default/files/2023-05/ACC_LegalTech-Report23_Final.pdf) (accessed: 24.02.2025).
2. *Rissland E.L.* Artificial intelligence and law: stepping stones to a model of legal reasoning // *The Yale Law Journal*. – 1990. – Vol. 99 (8). – P. 1957–1981. – DOI: 10.2307/796679.
3. *Соловьев И.П.* Изучение логического программирования на примере задач искусственного интеллекта // Компьютерные инструменты в образовании. – 1998. – № 6. – URL: <https://cyberleninka.ru/article/n/izuchenie-logicheskogo-programmirovaniya-na-primere-zadach-iskusstvennogo-intellekta> (дата обращения: 03.03.2025).
4. *Jaffar J., Lassez J.-L., Maher M.J.* Comments on “general failure of logic programs” // *The Journal of Logic Programming*. – 1986. – Vol. 3 (2). – P. 115–118. – DOI: 10.1016/0743-1066(86)90018-X.
5. *Leith P.* The rise and fall of the legal expert system† // *International Review of Law, Computers & Technology*. – 2016. – Vol. 30 (3). – P. 94–106. – DOI: 10.1080/13600869.2016.1232465.
6. *Loevinger L.* Jurimetrics: the methodology of legal inquiry // *Law and Contemporary Problems*. – 1963. – Vol. 28. – P. 5–35. – URL: <https://scholarship.law.duke.edu/lcp/vol28/iss1/2> (accessed: 03.03.2025).
7. *Genesereth M.* Computational law: the cop in the backseat. White Paper. – CodeX: The Stanford Center for Legal, 2015.
8. *Eken G.* Using natural language processing for automated construction contract review during risk assessment at the bidding stage. Ph.D. Doctoral Program. – Middle East Technical University, 2022.
9. Construction contract risk identification based on knowledge-augmented language models / S. Wong, C. Zheng, X. Su, Y. Tang // *Computers in Industry*. – 2024. – Vol. 157–158. – P. 104082. – DOI: 10.1016/j.compind.2024.104082.
10. *Moon S., Lee G., Chi S.* Automated system for construction specification review using natural language processing // *Advanced Engineering Informatics*. – 2022. – Vol. 51. – P. 101495. – DOI: 10.1016/j.aei.2021.101495.
11. *Shuai B., Caldas C.H.* A case-based rag methodology to analyze contract risks for construction projects // *SSRN Electronic Journal*. – 2024. – DOI: 10.2139/ssrn.4946907.
12. *Hassan F. ul, Le T.* Computer-assisted separation of design-build contract requirements to support subcontract drafting // *Automation in Construction*. – 2021. – Vol. 122. – P. 103479. – DOI: 10.1016/j.autcon.2020.103479.
13. *Kanapala A., Pal S., Pamula R.* Text summarization from legal documents: a survey // *Artificial Intelligence Review*. – 2019. – Vol. 51 (3). – P. 371–402. – DOI: 10.1007/s10462-017-9566-2.
14. *Dyevre A.* Text-mining for lawyers: how machine learning techniques can advance our understanding of legal discourse // *SSRN Electronic Journal*. – 2020. – DOI: 10.2139/ssrn.3734430.
15. *Bansal N., Sharma A., Singh R.K.* A review on the application of deep learning in legal domain // *IFIP International Conference on Artificial Intelligence Applications and Innovations*. – 2019. – Vol. 559. – P. 374–381. – DOI: 10.1007/978-3-030-19823-7\_31.
16. Method and system for suggesting revisions to an electronic document: Patent US2019/0303435 A1. – URL: <https://patents.google.com/patent/US20190303435A1/en> (accessed: 03.03.2025).
17. Systems, methods and software for processing phrases and clauses in legal documents: Patent US8788523B2. – URL: <https://patents.google.com/patent/US8788523B2/en> (accessed: 03.03.2025).
18. Systems, methods, and computer program products for slot normalization of text data: Patent US10614157B1. – URL: <https://patents.google.com/patent/US10614157B1/en> (accessed: 03.03.2025).
19. Eisenberg M. Mistake in contract law // *California Law Review*. – 2003. – Vol. 91 (6). – P. 1573. – DOI: 10.2307/3481399.
20. *Спесивцев А.В.* Нечетко-возможностный подход к формализации и использованию экспертных знаний для оценивания состояний сложных объектов // *Известия вузов. Приборостроение*. – 2020. – Т. 63, № 11. – С. 985–994. – DOI: 10.17586/0021-3454-2020-63-11-985-994.
21. Результаты опроса по квалификации юридической значимости ошибок в лицензионных договорах на ПО. – URL: [https://docs.google.com/spreadsheets/d/1TzE8BIL8BT9TpURq61OuCa3RZOyeYISqQ\\_7ZZ2CVUno/edit?resourcekey=&gid=97149320#gid=97149320](https://docs.google.com/spreadsheets/d/1TzE8BIL8BT9TpURq61OuCa3RZOyeYISqQ_7ZZ2CVUno/edit?resourcekey=&gid=97149320#gid=97149320) (дата обращения: 14.02.2024).
22. CUAD: an expert-annotated NLP dataset for legal contract review / D. Hendrycks, C. Burns, A. Chen, S. Ball. – 2021. – DOI: 10.48550/arxiv.2103.06268.
23. *Park J., Cardie C.* A corpus of erulemaking user comments for measuring evaluability of arguments // *Proceedings of the Eleventh International Conference on Language Resources and Eva-*

uation (LREC 2018), Miyazaki, Japan. – European Language Resources Association (ELRA), 2018. – P. 1623–1628.

24. Unsupervised alignment of privacy policies using hidden Markov models / R. Ramanath, F. Liu, N. Sadeh, N.A. Smith // Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. – Association for Computational Linguistics, 2014. – Vol. 2. – P. 605–610. – DOI: 10.3115/v1/P14-2099.

25. Question answering for privacy policies: combining computational and legal perspectives / A. Ravichander, A.W. Black, S. Wilson, T. Norton, N. Sadeh // Conference on Empirical Methods in Natural Language Processing, Hong Kong, November 2019. – DOI: 10.48550/arxiv.1911.00841.

26. *Koreeda Y., Manning C.D.* ContractNLI: a dataset for document-level natural language inference for contracts // Conference on Empirical Methods in Natural Language Processing. – 2021. – DOI: 10.48550/arxiv.2110.01799.

27. The black box problem revisited. Real and imaginary challenges for automated legal decision making / B. Brożek, M. Furman, M. Jakubiec, B. Kucharzyk // Artificial Intelligence and Law. – 2024. – Vol. 32 (2). – P. 427–440. – DOI: 10.1007/s10506-023-09356-9.

28. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions / L. Huang, W. Yu, W. Ma, W. Zhong, Z. Feng, H. Wang, Q. Chen, W. Peng, X. Feng, B. Qin, T. Liu // ACM Transactions on Information Systems. – 2025. – Vol. 43 (2). – P. 1–55. – DOI: 10.1145/3703155.

29. Alignment faking in large language models / R. Greenblatt, C. Denison, B. Wright, et al. – 2024. – DOI: 10.48550/arxiv.2412.14093.

30. Understanding the role of temperature in diverse question generation by GPT-4 / A. Agarwal, K. Mittal, A. Doyle, P. Sridhar, Z. Wan, J.A. Doughty, J. Savelka, M. Sakr // Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024). Vol. 2. – Association for Computing Machinery, 2024. – P. 1550–1551. – DOI: 10.1145/3626253.3635608.

*Котов Александр Александрович*, младший научный сотрудник, аспирант Федерального государственного бюджетного учреждения науки «Санкт-Петербургский федеральный исследовательский центр Российской академии наук». E-mail: alexanderkotovspb@gmail.com

*Kotov Alexander A.*, junior researcher, postgraduate student, St. Petersburg Federal Research Center of the Russian Academy of Sciences. E-mail: alexanderkotovspb@gmail.com

DOI: 10.17212/2782-2001-2025-1-27-48

### ***Problems of Developing Automated Contract Verification Systems\****

*A.A. KOTOV*

*St. Petersburg Federal Research Center of the Russian Academy of Sciences, 39, 14th line of Vasilyevsky Island, St. Petersburg, 199178, Russian Federation*

*alexanderkotovspb@gmail.com*

#### **Abstract**

This paper is dedicated to the development of an automated contract verification system and the challenges encountered during its implementation. The relevance of the research lies in addressing the tasks of optimizing the work of contract review experts.

The study identifies key problems in this area, including: 1) the lack of a standardized risk assessment system, which complicates decision-making for decision-makers; 2) the lack of data for training the system, limiting its ability to effectively detect errors and correctly interpret legal terms; 3) insufficient control over the process when using large language models, which, despite their powerful analytical capabilities, do not always guarantee the accuracy and consistency of conclusions; 4) lack of consideration for context and semantic variability in legal texts, which complicates accurate classification and error detection, especially in cases where the same legal terms or phrases may have different meanings depending the context.

---

\* Received 19 November 2024.

As a result of the research, the following findings were obtained: 1) an original risk classification system for legal errors based on the fuzzy set theory; 2) algorithms for the operation of the legal document review system; 3) methods for generating data to fine-tune BERT models and adjust GPT models.

Furthermore, the article proposes a combined approach using GPT and BERT models (poly-model system) to optimize the process: the GPT model processes the text and identifies errors based on predefined prompts, while the BERT model is used to assign risk levels to the errors and control the results of the GPT model, ensuring a more reliable and consistent analysis.

The article also emphasizes the need to create open datasets of contracts based on the Russian law. Creating such resources will not only improve the quality of training for artificial intelligence systems but also enhance the accessibility and transparency of information for researchers and practitioners in the field of law. This will be an important step toward improving automated legal document analysis and increasing overall legal confidence when working with them.

**Keywords:** automated contract verification, legal metrics, large language models, legal-tech, computational law, data analysis, artificial intelligence, risk assessment, standardization

## REFERENCES

1. *2023 Legal Technology Report: for In-house Legal Professionals*. Association of Corporate Counsel, 2023. Available at: [https://www.acc.com/sites/default/files/2023-05/ACC\\_LegalTech-Report23\\_Final.pdf](https://www.acc.com/sites/default/files/2023-05/ACC_LegalTech-Report23_Final.pdf) (accessed 24.02.2025).
2. Rissland E.L. Artificial intelligence and law: stepping stones to a model of legal reasoning. *The Yale Law Journal*, 1990, vol. 99 (8), pp. 1957–1981. DOI: 10.2307/796679.
3. Solov'ev P. Izuchenie logicheskogo programmirovaniya na primere zadach iskusstvennogo intellekta [Study of logical programming on the example of artificial intelligence problems]. *Komp'yuternye instrumenty v obrazovanii = Computer Tools in Education*, 1998, no. 6. Available at: <https://cyberleninka.ru/article/n/izuchenie-logicheskogo-programmirovaniya-na-primere-zadach-iskusstvennogo-intellekta> (accessed 03.03.2025).
4. Jaffar J., Lassez J.-L., Maher M.J. Comments on “general failure of logic programs”. *The Journal of Logic Programming*, 1986, vol. 3 (2), pp. 115–118. DOI: 10.1016/0743-1066(86)90018-X.
5. Leith P. The rise and fall of the legal expert system†. *International Review of Law, Computers & Technology*, 2016, vol. 30 (3), pp. 94–106. DOI: 10.1080/13600869.2016.1232465.
6. Loevinger L. Jurimetrics: the methodology of legal inquiry. *Law and Contemporary Problems*, 1963, vol. 28, pp. 5–35. Available at: <https://scholarship.law.duke.edu/lcp/vol28/iss1/2> (accessed 03.03.2025).
7. Genesereth M. *Computational law: the cop in the backseat*. White Paper. CodeX: The Stanford Center for Legal, 2015.
8. Eken G. *Using natural language processing for automated construction contract review during risk assessment at the bidding stage*. Ph.D. Doctoral Program. Middle East Technical University, 2022.
9. Wong S., Zheng C., Su X., Tang Y. Construction contract risk identification based on knowledge-augmented language models. *Computers in Industry*, 2024, vol. 157–158, p. 104082. DOI: 10.1016/j.compind.2024.104082.
10. Moon S., Lee G., Chi S. Automated system for construction specification review using natural language processing. *Advanced Engineering Informatics*, 2022, vol. 51, p. 101495. DOI: 10.1016/j.aei.2021.101495.
11. Shuai B., Caldas C.H. A case-based rag methodology to analyze contract risks for construction projects. *SSRN Electronic Journal*, 2024. DOI: 10.2139/ssrn.4946907.
12. Hassan F. ul, Le T. Computer-assisted separation of design-build contract requirements to support subcontract drafting. *Automation in Construction*, 2021, vol. 122, p. 103479. DOI: 10.1016/j.autcon.2020.103479.
13. Kanapala A., Pal S., Pamula R. Text summarization from legal documents: a survey. *Artificial Intelligence Review*, 2019, vol. 51 (3), pp. 371–402. DOI: 10.1007/s10462-017-9566-2.
14. Dyeve A. Text-mining for lawyers: how machine learning techniques can advance our understanding of legal discourse. *SSRN Electronic Journal*, 2020. DOI: 10.2139/ssrn.3734430.
15. Bansal N., Sharma A., Singh R.K. A review on the application of deep learning in legal domain. *IFIP International Conference on Artificial Intelligence Applications and Innovations*, 2019, vol. 559, pp. 374–381. DOI: 10.1007/978-3-030-19823-7\_31.
16. *Method and system for suggesting revisions to an electronic document*. Patent US 2019/0303435 A1. Available at: <https://patents.google.com/patent/US20190303435A1/en> (accessed 03.03.2025).

17. *Systems, methods and software for processing phrases and clauses in legal documents*. Patent US8788523B2. Available at: <https://patents.google.com/patent/US8788523B2/en> (accessed 03.03.2025).
18. *Systems, methods, and computer program products for slot normalization of text data*. Patent US10614157B1. Available at: <https://patents.google.com/patent/US10614157B1/en> (accessed 03.03.2025).
19. Eisenberg M. Mistake in contract law. *California Law Review*, 2003, vol. 91 (6), p. 1573. DOI: 10.2307/3481399.
20. Spesivtsev A.V. Nechetko-vozmozhnostnyi podkhod k formalizatsii i ispol'zovaniyu ekspertnykh znaniy dlya otsenivaniya sostoyanii slozhnykh ob"ektov [Fuzzy-probabilistic approach to formalizing and using expert knowledge to evaluate complex objects states]. *Izvestiya vuzov. Priborostroenie = Journal of Instrument Engineering*, 2020, vol. 63, no. 11, pp. 985–994. DOI: 10.17586/0021-3454-2020-63-11-985-994.
21. *Survey results on the qualification of legal significance of errors in software license agreements*. Available at: [https://docs.google.com/spreadsheets/d/1TzE8BIL8BT9TpURq61Ouca3RZOyeYISqQ\\_7ZZ2CVUno/edit?](https://docs.google.com/spreadsheets/d/1TzE8BIL8BT9TpURq61Ouca3RZOyeYISqQ_7ZZ2CVUno/edit?) (accessed 14.02.2024).
22. Hendrycks D., Burns C., Chen A., Ball S. *CUAD: an expert-annotated NLP dataset for legal contract review*. 2021. DOI: 10.48550/arxiv.2103.06268.
23. Park J., Cardie C. A corpus of erulemaking user comments for measuring evaluability of arguments. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA), 2018, pp. 1623–1628.
24. Ramanath R., Liu F., Sadeh N., Smith N.A. Unsupervised alignment of privacy policies using hidden Markov models. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2014, vol. 2, pp. 605–610. DOI: 10.3115/v1/P14-2099.
25. Ravichander A., Black A.W., Wilson S., Norton T., Sadeh N. Question answering for privacy policies: combining computational and legal perspectives. *Conference on Empirical Methods in Natural Language Processing*, Hong Kong, November 2019. DOI: 10.48550/arxiv.1911.00841.
26. Koreeda Y., Manning C.D. ContractNLI: a dataset for document-level natural language inference for contracts. *Conference on Empirical Methods in Natural Language Processing*, 2021. DOI: 10.48550/arxiv.2110.01799.
27. Brožek B., Furman M., Jakubiec M., Kucharzyk B. The black box problem revisited. Real and imaginary challenges for automated legal decision making. *Artificial Intelligence and Law*, 2024, vol. 32 (2), pp. 427–440. DOI: 10.1007/s10506-023-09356-9.
28. Huang L., Yu W., Ma W., Zhong W., Feng Z., Wang H., Chen Q., Peng W., Feng X., Qin B., Liu T. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 2025, vol. 43 (2), pp. 1–55. DOI: 10.1145/3703155.
29. Greenblatt R., Denison C., Wright B., et al. *Alignment faking in large language models*. 2024. DOI: 10.48550/arxiv.2412.14093.
30. Agarwal A., Mittal K., Doyle A., Sridhar P., Wan Z., Doughty J.A., Savelka J., Sakr M. Understanding the role of temperature in diverse question generation by GPT-4. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024)*. Vol. 2. Association for Computing Machinery, 2024, pp. 1550–1551. DOI: 10.1145/3626253.3635608.

Для цитирования:

Котов А.А. Проблемы разработки систем автоматизированной проверки договоров // Системы анализа и обработки данных. – 2025. – № 1 (97). – С. 27–48. – DOI: 10.17212/2782-2001-2025-1-27-48.

For citation:

Kotov A.A. Problemy razrabotki sistem avtomatizirovannoi proverki dogovorov [Problems of developing automated contract verification systems]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2025, no. 1 (97), pp. 27–48. DOI: 10.17212/2782-2001-2025-1-27-48.