Автономная некоммерческая организация высшего образования
«Университет Иннополис»

АННОТАЦИЯ
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ
(МАГИСТЕРСКУЮ ДИССЕРТАЦИЮ)
ПО НАПРАВЛЕНИЮ ПОДГОТОВКИ
09.04.01 – «ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА»

НАПРАВЛЕННОСТЬ (ПРОФИЛЬ) ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ
«АНАЛИЗ ДАННЫХ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ»

Тема **Обработка естественного языка для аудита соответствия нормативным требованиям. Автоматизированная проверка соответствия соглашений об обработке данных на основе обработки естественного языка**

Выполнил **Оконича Озиома Ненубари**

подпись

Иннополис, Innopolis, 2024

Оглавление

Аннотация

В современном мире соблюдение нормативных актов, касающихся защиты данных, таких как GDPR, играет центральную роль для организаций. Другой важной проблемой, выявленной в ходе анализа, является тот факт, что соблюдение требований затруднено из-за того, что юридические документы часто бывают сложными, а правила постоянно меняются. Цель данной работы - описать способы, с помощью которых НЛП помогает обеспечить соответствие GDPR без особых усилий путем автоматического сканирования на предмет соответствия, оценки политик конфиденциальности и повышения уровня прозрачности.

Работа не ограничивается изучением применения NLP для работы с политиками конфиденциальности и облегчения лучшего понимания совместного использования данных третьими лицами, но также предполагает проведение экспериментов для оценки разницы между несколькими моделями NLP. Они внедряют и выполняют модели, чтобы выделить ту, которая работает лучше всего, основываясь на эффективности и скорости, с которой она автоматизирует процесс проверки соответствия и анализа политики конфиденциальности.

Кроме того, некоторые вопросы, затронутые в исследовании, относятся к автоматизированным решениям GDPR, таким как генерация машиночитаемых моделей, которые делают процесс оценки соответствия более эффективным.

Таким образом, в данной работе подчеркивается важность использования НЛП для того, чтобы помочь организациям преодолеть трудности,

связанные с соблюдением GDPR, и создать дорожную карту для перехода к более клиентоориентированному режиму защиты данных. В этой связи, сравнивая проведенные эксперименты и показывая эффективность лучшей модели, она помогает повысить эффективность мер, принимаемых для обеспечения соответствия, и способствует защите прав личности в киберпространстве.

Глава 1

Введение

Учитывая тот факт, что в современной цифровой среде люди производят множество персональных данных и обмениваются ими, защита частной жизни становится все более актуальной. Ведущую роль в этом играет Общий регламент по защите данных (GDPR) - соответствующий правовой инструмент и свод законов, принятых ЕС для смягчения обработки персональных данных и защиты прав граждан на неприкосновенность частной жизни и защиту данных. В целом GDPR решает проблему защиты доверия, налагая строгие законодательные требования на организации, обрабатывающие персональные данные, и регулируя деятельность по обработке данных на основе принципов прозрачности, подотчетности и согласия.

Ниже приведены некоторые определения ключевых слов, которые будут встречаться в диссертации:

- **Обработка естественного языка (NLP):** отрасль информатики, которая занимается тем, что заставляет компьютерные приложения понимать и анализировать письменный или устный человеческий язык [1].

- **Соглашения об обработке данных (DPAs):** Юридически обязывающие документы, заключаемые между контроллером и процессором в письменной или электронной форме [2].

- **GDPR (General Data Protection Regulation):** Всеобъемлющая законодательная база, принятая Европейским союзом (ЕС) для регулирования обработки персональных данных и защиты прав граждан на конфиденциальность и защиту данных [3].

- **Приватность:** Предоставление пользователям возможности самим принимать решения о том, кто может обрабатывать их данные и с какой целью [4].

- **Соблюдение требований:** Выполнение требований по надлежащей обработке персональных данных в соответствии с законом [5].

- **Метод проверки соответствия:** Процесс обеспечения соответствия организации требованиям GDPR.

- **Аудит соответствия:** Систематическая и независимая оценка соответствия организации требованиям GDPR [6].

- **Text Mining:** Получение высококачественной информации из текста [7].

- **Трансферное обучение:** Техника машинного обучения (ML), при которой знания, полученные в ходе выполнения задачи, повторно используются для повышения производительности при выполнении смежной задачи [8].

- **BERT:** Архитектура трансформатора "только для кодирования"и базовая модель в НЛП [9].

1.1 Background

В этом документе раздел "История вопроса"содержит информацию о том, почему появился Общий регламент по защите данных (GDPR). Его корни уходят к некоторым проблемам, возникшим в середине двадцатого века, связанным с управлением личной информацией. С появлением Интернета как одной из коммуникационных технологий вышеупомянутые проблемы были реализованы в полной мере. Сбор, использование и обмен персональными данными вызывали растущую озабоченность общественности в связи с утечками данных, которые случались практически во всех организациях [10].

В связи с ростом использования интернет-приложений и различных онлайн-сервисов увеличился и объем обрабатываемых персональных данных. Это, безусловно, вызвало серьезную озабоченность вопросами конфиденциальности и безопасности. Были случаи утечки конфиденциальной информации и личных идентификационных данных в чужие руки, использования информации не по назначению, а также случаи, когда нам не сообщали, как используется наша информация [11]. В данной статье мы попытаемся выделить и объяснить эти проблемы, а также показать, как GDPR стремится преодолеть их, давая указания организациям при работе с персональными данными.

1.2 Значение НЛП в обеспечении соответствия GDPR

Раздел "Понимание роли NLP в GDPR"определяет роль обработки естественного языка в обеспечении соблюдения организациями политик GDPR. Он начинается с понимания того, что юридические тексты, такие как политики конфиденциальности, которые важны для каждого пользователя Интернета, вряд ли могут быть легко понятны, поскольку содержат множество терминов, чуждых обычному человеку. Как правило, проверка на соответствие требованиям уже давно осуществляется вручную, и это довольно трудоемко, а также сопряжено с вероятностью ошибки.

НЛП удовлетворяет эту потребность за счет использования интеллектуальных алгоритмов, которые помогают анализировать и понимать человеческий язык. Таким образом, используя методы НЛП, организации могут повысить эффективность своей деятельности, связанной с анализом и пониманием юридических текстов. Речь идет о самых разных задачах - от анализа содержащихся в них положений до извлечения определенной информации и определения соответствия различным аспектам, связанным с GDPR. Применение НЛП полезно не только для повышения эффективности процесса обеспечения соответствия, но и для повышения точности и согласованности при оценке соответствия организаций нормативным требованиям [12], [13].

Кроме того, с помощью НЛП предприятия становятся более способными реагировать на изменения в законодательстве по защите конфиденциальности. При работе с моделями NLP можно удовлетворять юридические потребности, а по мере изменения юридических трактовок и добавления новых требований модели можно обучать заново. Это гарантирует, что ор-

ганизации будут соблюдать требования GDPR на протяжении всего времени и в то же время будут информированы о вопросах, связанных с защитой данных и правилами конфиденциальности. Наконец, НЛП позволяет организациям более эффективно решать проблемы, связанные с соблюдением GDPR, чтобы сохранить права людей в контексте цифровой трансформации.

1.3 Постановка проблемы

Вопросы, описанные в этом разделе, связаны с проблемами, с которыми сталкиваются организации при попытке соблюдения положений GDPR и расшифровки неоднозначных правовых текстов. В нем освещаются такие типичные условия, как как понять, что на самом деле говорят правила, законы в разработке и как донести концепцию политики конфиденциальности до неюридической аудитории? Цель этой части диссертации - объяснить эти проблемы, чтобы к концу работы мы могли попытаться ответить на следующие вопросы.

Прежде всего, необходимо отметить, что работа с правовыми нормами, предусмотренными в рамках GDPR, является весьма проблематичной для многих организаций, работающих в условиях цифровой среды. Одна из них - сложность, возникающая в связи с необходимостью осмыслить язык, используемый в законодательстве, который может быть трудно понять без юридического образования. Именно в связи с этим GDPR включает в себя множество правил и положений [14], которые иногда трудно понять, не говоря уже о том, чтобы выполнить.

Однако GDPR не является чем-то застывшим во времени; скорее, когда он вводится в действие, он меняется со временем, постоянно вносятся по-

правки, уточнения и новые правила. Это придает процессу еще один уровень динамизма, поскольку организация должна быть информирована о соответствующих изменениях и корректировать свои действия в соответствии с ними [15].

Кроме того, внедрение принципов защиты данных и политики конфиденциальности пользователей также представляет собой большую проблему с точки зрения того, как эти положения доносятся до пользователей. В большинстве письменных документов, особенно в политике конфиденциальности, используются юридические термины, которые многим людям трудно понять, и это заставляет таких людей не знать о своих правах и о том, как будут использоваться имеющиеся у них данные.

Эти проблемы объясняют, почему гибкие, нестандартные, простые в использовании решения, способные создавать положения GDPR простым языком и без особых усилий, могут значительно облегчить понимание юридических текстов, изложенных в соответствии с GDPR, а также помочь повысить прозрачность практики обработки данных. Решение этих задач может помочь организациям защитить права людей на неприкосновенность частной жизни и защиту данных и должным образом выполнить требования законодательства.

1.4 Вопросы и цели исследования

В следующей части статьи мы представляем вопросы исследования, которыми мы руководствуемся, и цели, которые мы ставим перед собой. Эти вопросы помогут нам перейти к более глубокому анализу и рекомендациям в области соответствия GDPR и обработки естественного языка (NLP).

- Насколько эффективны модели НЛП, включая GPT-3, для автоматизации выявления проблем соответствия GDPR в политике конфиденциальности данных организации?
- Каковы ограничения текущих технологий НЛП в интерпретации и обеспечении соответствия GDPR, и как эти ограничения могут быть устранены?
- Какую роль могут сыграть инструменты на базе НЛП в поддержке специалистов по соблюдению нормативных требований и юристов в обеспечении соответствия GDPR?

Задачи ясны: глубже понять, насколько эффективно НЛП помогает улучшить соответствие нормативным требованиям, и разработать больше эффективных инструментов для руководства организаций по соблюдению законов о защите данных. Таким образом, цель - провести глубокое исследование и, по возможности, расследование, чтобы помочь людям и организациям придерживаться правил, установленных GDPR, расширив базу знаний в области конфиденциальности данных и соответствия.

1.5 Структура диссертации

В этой части работы мы постараемся объяснить, как представлена вся диссертация и что рассматривается в каждом из разделов. Она дает общее представление о том, чего следует ожидать.

Во-первых, это введение, поскольку оно дает предпосылку для последующих действий, излагая общую проблему и ее значение. Наконец, мы

переходим к обзору литературы, чтобы выяснить, что уже было сделано другими в этой области. Далее следует метод, где мы рассказываем о том, как мы проводили исследование после обзора литературы и теоретической формулировки рассматриваемых вопросов.

Экспериментальные результаты, к которым мы приступили, приведены в следующем разделе работы. Наконец, в заключении представлены выводы, в которых подчеркиваются итоги исследования и его последствия.

Глава 2

Обзор литературы

В последние несколько лет тема соответствия нормативным требованиям и, в частности, в связи с GDPR как примером законодательства о защите данных, не может не привлекать внимание к кажущимся *"революционными"* изменениям, которые были вызваны включением возможностей NLP [12], [16], [17]. В результате этой эволюции НЛП стало перспективной областью для автоматизации различных процессов, связанных с проверкой данных и политикой конфиденциальности на основе GDPR и других законов о защите данных [18]—[20].

Объединение возможностей NLP с деятельностью по соблюдению правовых норм - это революция в понимании и работе с нормативными требованиями и юридическими документами. Как правило, соблюдение требований по защите данных отнимало много времени и средств, что влекло за собой бумажное просеивание и анализ огромных юридических текстов [21]—[23]. Хотя информации об использовании NLP в работе с нормативно-правовым соответствием не так много, его применение принесло эффективность и масштабируемость в использовании машинного обучения и понимания есте-

ственного языка в организации для работы с нормативно-правовым соответствием [24]—[26].

2.1 Обзор

Таким образом, этот обширный обзор литературы направлен на обсуждение ключевых публикаций в этой области и освещает различные методологии, модели и подходы, которые опираются на НЛП для анализа сложности соответствия GDPR и других юридических документов. Рассматривая специфику законодательства о защите данных и конфиденциальности, специалисты по защите данных и ученые-первопроходцы разработали совершенно новые подходы, которые касаются трудностей в соответствующих процедурах соответствия и разъяснения практики использования данных в таких документах, как политика конфиденциальности, соглашения об обработке данных и нормативные требования [27]—[29].

Ключевые области, на которых сосредоточен данный обзор литературы, включают:

- **Проверка соответствия:** Существуют хорошо продуманные и продвинутые модели и фреймворки, которые описывают, как можно автоматизировать проверку соответствия требованиям GDPR. Эти модели действуют как юридические консультанты, которые используют стратегии NLP для анализа юридических текстов и определения соответствия организации требованиям правовых прецедентов [3], [30], [31].

- **Анализ политики конфиденциальности:** Корпорации, а также такие инструменты, как PrivaSeer, стали мощными NLP-решениями для

анализа политик и значительно помогли в крупномасштабном сборе данных для извлечения, а также классификации [32]—[34]. Эти представления могут быть использованы для повышения или улучшения текущего состояния прозрачности, а также для улучшения способности пользователей делать правильный выбор в отношении их права на конфиденциальность и защиту своих данных.

- **Семантическая аннотация:** Предыдущие попытки индексирования и аннотирования юридических текстов с помощью семантики позволили добиться улучшения работы поиска, а также извлечения информации [35]—[37]. Добавляя метаданные и семантические теги в юридические тексты, ученые создали возможности для практической работы в будущем, чтобы улучшить доступность и понимание правовых норм.

Таким образом, именно в рамках использования этих пионеров в области НЛП данный обзор литературы направлен на раскрытие возможностей методологии НЛП в контексте GDPR и защиты данных. Таким образом, интегрируя различные методологии и подходы, ученые пытаются усилить инновации, повысить роль прозрачности методов и помочь организациям эффективно справляться с вызовами постоянно меняющейся правовой среды [21], [28], [38].

2.2 Анализ политики приватности: Сокращение легализации для повышения прозрачности

Политика конфиденциальности - это один из важнейших юридических инструментов, предоставляющих информацию об управлении данными ор-

ганизации от многочисленных пользователей. Тем не менее, такие политики написаны на юридическом жаргоне, и поэтому пользователи с трудом понимают, что в них говорится. В ответ на эту проблему ученые стали применять уникальные методы, основанные на НЛП, для автоматизации оценки политик конфиденциальности с целью повышения прозрачности [25], [39], [40].

[25] представляет инновационный проект, известный как PrivaSeer, который сосредоточен на проблеме сложности политик конфиденциальности и возможных способах помочь пользователям легче и лучше понять их. Около миллиона политик конфиденциальности англоязычных сайтов включены в набор данных, названный PrivaSeer, который является самым большим набором данных такого рода. Накопление массива осуществляется с помощью сложных процессов веб-ползания и фильтрации, которые составляют конвейер создания и гарантируют полноту и релевантность массива. Поскольку он предлагает исследователям список ссылок на политики конфиденциальности, которые содержат как сильные, так и слабые стороны для сравнения, анализ информации с помощью PrivaSeer и определение областей для улучшения формулировки политики конфиденциальности могут быть эффективно достигнуты.

Аналогичным образом [33] рассматривает идентификацию и категоризацию сторонних организаций из политик конфиденциальности, связанных с мобильными приложениями. В связи с этим в исследовании признается необходимость уточнения получателей данных в связи с восприятием пользователями конфиденциальности приложений. Используя современные модели распознавания именованных сущностей (NER), исследование стремится улучшить осведомленность пользователей о сторонних сущностях, присутствующих в политиках конфиденциальности. Благодаря использова-

нию аннотаций и подробных моделей, данное исследование направлено на создание эффективных и действенных инструментов для извлечения надлежащего набора информации из политик конфиденциальности, что позволит заинтересованным сторонам принимать рациональные решения по защите своих данных.

2.3 Проверка соответствия GDPR: Автоматизация соблюдения нормативных требований

GDPR по-прежнему вызывает серьезную озабоченность у организаций, особенно с точки зрения уровня соответствия законам о защите данных. В связи со сложностью и активностью мандатов GDPR концепция проверки должна была развиваться, и одним из потенциальных решений является обработка естественного языка [20], [31], [41].

[20] следует инженерному подходу, основанному на модели, чтобы предложить концептуальную модель, которая может быть проанализирована машиной для захвата характеристик DPA в соответствии с GDPR. Таким образом, разрабатывая конкретные критерии для оценки соответствия DPA и предлагая автоматизированный инструмент для проверки соответствия, исследование направлено на облегчение давления на контролеров и обработчиков данных. Таким образом, DPA разработали концептуальную модель, которая является основой для оценки соответствия организаций GDPR и поможет в раннем обнаружении и предотвращении проблем с соответствием.

В этой же статье [18] обсуждается проект SPECIAL H2020, в рамках

которого будут предоставлены инструменты, позволяющие организациям проводить автоматические проверки соответствия. В основе проекта лежит ключевая тема - разработка языка политик, способного выражать согласие, бизнес-политику и нормативные обязательства таким образом, чтобы машина могла их понять. Проект предусматривает два принципиально разных варианта автоматизированной проверки соответствия, что позволяет предложить организациям гибкие возможности для проверки соответствия GDPR. SPECIAL H2020 - это инновационный проект, включающий автоматизированные NLP-решения, направленные на повышение соответствия требованиям и стандартам в различных областях.

2.4 Семантическая аннотация и извлечение юридических метаданных: Создание превосходного юридического обоснования

Семантическая аннотация юридических текстов является важнейшей предпосылкой для реализации таких элементов, как намерение поисковика, поддержка процесса юридической аргументации и интерпретационной масштабируемости. Эти заметные успехи в данной области делают данную работу [42] важным источником информации, поскольку в ней представлен смешанный подход к извлечению семантических юридических метаданных с помощью обработки естественного языка (NLP)[17], [28], [43]. Таким образом, предлагая единую концептуальную модель типов семантических метаданных, связанных с анализом правовых требований, и список правил автоматического извлечения на основе синтаксического анализа конституен-

тов и зависимостей, исследование закладывает основу для систематического анализа правовых положений. Таким образом, так называемая концептуальная модель предполагает разграничение понятий на уровне фраз и на уровне высказываний, что завершает создание основы для семантики правовых положений. Исследование, проведенное с помощью токенизации, разбиения предложений, POS-тегирования, NER, а также конституентного и зависимого парсеров, показывает тонкое извлечение семантических правовых метаданных. Таким образом, проверка правил извлечения дала только положительные результаты, в то время как точность варьируется от $87,4\%$ до $97,2\%$, recall - от $85,5\%$ до $94,9\%$, что говорит об эффективности предложенного подхода в достижении скудной юридической семантики [43].

Аналогично, [17], [44] предлагает работу о прагматическом подходе к семантическому аннотированию и методологии проектирования CLAL в контексте XML с надлежащей формализацией. В этом отношении исследование относится к улучшению правовых положений путем семантического аннотирования текстов, что просто добавляет интерпретационно-нейтральную информацию в качестве метаданных к правовым текстам и таким образом улучшает поиск, а также более эффективный правовой анализ положений. CLAL играет роль языка аннотирования и является общим языком, используемым для аннотирования юридических текстов, и включает в себя достаточно большое количество семантических характеристик, необходимых для правового рассуждения и понимания. Это координируется с помощью мер межаннотаторского согласия, которые используются для подтверждения надежности и безопасности языка аннотации в процессе аннотирования [17]. Данный аннотированный корпус в сочетании с набросками схемы CLAL может быть полезен для развития дальнейших исследований и практического

использования, направленного на поиск и анализ юридических текстов с расширенными критериями современного изучения.

2.5 Трудности использования НЛП для обеспечения соответствия GDPR

Тем не менее, следует отметить некоторые проблемы, доказывающие необходимость продолжения исследований и разработок в области NLP для GDPR, а также анализа юридических текстов в целом. Эти вопросы должны быть решены в дальнейшем с целью предоставления специализированных решений, которые являются надежными и могут легко решать различные аспекты, связанные с регулированием и юридической документацией [12], [20], [45].

1. Ограничения набора данных: Одной из проблем является тот факт, что для достижения наилучших результатов эти модели должны обучаться на более крупных и разнообразных наборах данных - отсюда и доступность. Существующие на данный момент наборы данных могут оказаться недостаточными для обеспечения необходимого объема и глубины охвата, чтобы отразить вариации использования юридического языка и связанные с этим ситуации соблюдения. Поэтому существует острая необходимость собрать соответствующие коллекции общих юридических работ, правовых кодексов и языковых различий [21], [28], [46].

2. Адаптируемость к развивающимся правовым рамкам: В случае с решениями по обеспечению соответствия на основе NLP правовые рамки - это GDPR и другие законы о защите данных, которые представляют

собой самую большую проблему из-за их изменчивости. Правовые тексты часто меняются, интерполируются, дополняются или просто переинтерпретируются, что требует от моделей соответствия этим изменениям, когда они происходят. Необходимы модели НЛП, достаточно гибкие и динамичные, чтобы следовать любым изменениям в правовых спецификациях на предмет соответствия требованиям [36], [47], [48].

3. Исследование новых архитектур НЛП: Несмотря на то, что существующие архитектуры НЛП показали хороший уровень достижений во многих областях, таких как проверка соответствия и юридический анализ текстов, исследователи все еще пытаются открыть новые архитектуры или предварительно обученные модели. Генеративные модели, включая GPT-3 и T5, можно использовать для создания естественных ответов, а также для понимания юридических документов на основе контекста. Было бы интересно опробовать эти нейронные архитектуры и включить их в задачи, связанные с соблюдением требований; здесь может быть гораздо больше потенциала для повышения точности и эффективности [21], [49], [50].

2.6 Будущие направления исследований

. На основе этих исследовательских пробелов и ограничений, упомянутых в настоящей работе, для дальнейшего развития NLP в области соответствия GDPR можно наметить следующие стратегии для будущих исследований:

1. Эксперименты с генеративными моделями: Метод использования генеративных моделей должен быть расширен в будущих исследованиях за счет современных генеративных моделей, таких как GPT-3 и T5. Эти моде-

ли обладают расширенными возможностями по обработке и генерации НЛ, которые особенно подходят для сложных приложений, таких как проверка соответствия, проверка юридических документов и подобных процессов [24], [25], [37].

2. Расширенные оценки: Очень важно проводить более комплексные исследования решений и приложений для обеспечения соответствия на основе НЛП на предмет того, насколько хорошо они работают, насколько они устойчивы и насколько легко их можно масштабировать. Система оценки должна содержать множество наборов данных, метрик и приложений, чтобы дать рекомендации по использованию подхода и информацию о его эффективности [3], [18], [45].

3. Междисциплинарные подходы: Еще одним возможным методом дальнейших исследований является фокусирование внимания как на юридических знаниях, включаемых в разработку решений по обеспечению соответствия, так и на методах НЛП для улучшения таких решений. Междисциплинарное сотрудничество со специалистами в области правоприменения, информационных технологий, комплаенса и НЛП может способствовать улучшению существующих решений путем создания инструментов, более осведомленных о специфических проблемах в сложной сфере комплаенса [27], [39], [43].

Если исследователи воспользуются этими проблемами и будущими направлениями развития НЛП для обеспечения соответствия GDPR, они смогут продвинуться в этой области в плане создания мощных, расширяемых и реагирующих моделей, которые помогут организациям приобрести навыки и инструменты, чтобы справиться с этими непредсказуемыми простыми рекомендательными услугами.

Подводя итог, можно сказать, что рассмотренная литература подтверждает важность НЛП для решения сложных проблем, связанных с регулированием GDPR, пониманием политики конфиденциальности и расшифровкой юридических документов. Необходимо провести дальнейшие исследования и эксперименты, чтобы заполнить этот пробел и повысить уровень техники в прикладных подходах для автоматизированной проверки соответствия и анализа юридических текстов.

Глава 3

Методология

3.1 Введение

В этой главе объясняется, как методы данной диссертации были использованы для обнаружения автоматизированного соблюдения Общего регламента по защите данных (GDPR) с помощью методов обработки естественного языка (NLP). Тот факт, что юридические тексты специфичны, а правовые нормы часто меняются, делает актуальным систематический подход к разработке, обучению и оценке моделей NLP. В данном исследовании использовались два различных набора данных: OPP-115 и ACL Coling. Они использовались для обучения и проверки эффективности нескольких моделей НЛП, включая SBERT, BERT и GPT, которые являются одними из самых мощных технологий обработки языков.

Первое в методологии - получение и предварительная обработка набора данных, на котором будут обучаться модели. Затем проводится тщательный процесс сопоставления категорий из набора данных OPP-115 с принципами GDPR 5. Это в дальнейшем поможет в обучении моделей для классификации

по нескольким меткам. Кроме того, мы анализируем различные гранулярности. Будет проведен анализ как на уровне предложений, так и на уровне всей политики, чтобы выяснить эффективность моделей в различных ситуациях.

Также в данной методологии будет дано исчерпывающее объяснение возможностей и ограничений выбранных моделей, используемых для выявления соответствия GDPR. В частности, будут подробно рассмотрены плюсы и минусы упомянутых моделей, поскольку каждая из них подчеркивает свои уникальные возможности и причину выбора. Поэтому я подчеркну их пригодность для решения поставленной задачи.

Суть методологии заключается в процессе обучения классификации по нескольким меткам с учетом принципов GDPR, подготовке отчетов о соответствии и метриках оценки, специально созданных для оценки производительности и точности моделей. Наконец, я объясняю методы тонкой настройки, использованные для повышения эффективности моделей.

Таким образом, в следующих разделах я опишу как технические, так и теоретические подходы, использованные для достижения целей исследования, указанных в главе 1.

3.2 Получение набора данных

Я использовал два набора данных, полученных из Полезные политики конфиденциальности [51]. Комбинация двух наборов данных позволяет получить более надежные результаты обучения и тестирования моделей. В то время как один из них позволяет целенаправленно обучать модели, поскольку в нем есть политики конфиденциальности, которые правильно аннотированы, другой набор данных ставит модели под сомнение. Не имея целей и

аннотаций, я смог проверить, насколько адаптируемы и обобщаемы модели.

Такой способ выбора наборов данных был стратегическим, и он гарантирует, что обученные модели не будут специализироваться только на проверке соответствия GDPR в специально определенных границах. Но они также способны хорошо работать и в других правовых контекстах.

3.2.1 OPP-115 Dataset

OPP-115, который происходит от Online Privacy Policy Project, представляет собой набор данных из 115 политик, собранных в 2016 году с различных веб-сайтов. Он послужил хорошей отправной точкой для обучения NLP-моделей на соответствие GDPR. Все политики в этом наборе данных подробно аннотированы по набору определенных категорий, которые соответствуют принципам, предусмотренным статьей 5 GDPR. Каждая политика была очищена и имеет красивую печать в формате html и соответствующую аннотацию в формате csv. Все они были просмотрены и помечены аннотациями, характеризующими конкретные элементы данных. Таким образом, этот набор данных очень хорошо подходит для обучения моделей пониманию нюансов политики конфиденциальности.

Если углубиться, то краткое описание категорий OPP-115 [52] выглядит следующим образом:

1. Сбор/использование первой стороной: что, почему и как информация собирается поставщиком услуг

2. Совместное использование/собирание третьей стороной: Что, почему и как информация передается третьим лицам или собирается ими.

3. Безопасность данных: Меры защиты пользовательской информации
4. Хранение данных: Как долго будет храниться информация о пользователе
5. Выбор/контроль пользователя: Возможности контроля, доступные пользователям
6. Доступ пользователей, редактирование и удаление: Если/как пользователи могут получить доступ, редактировать или удалять информацию
7. Изменение политики: Информирование пользователей о том, что информация о политике была изменена
8. Международная и специфическая аудитория: Практика, относящаяся к определенной группе пользователей
9. Прочее: Общий текст, контактная информация или методы, не входящие в другие категории.

Пример того, как выглядит аннотация, показан ниже. Текст, выделенный желтым цветом, показывает тип данных в категории, а текст, выделенный синим цветом, показывает извлеченное предложение, относящееся к категории.

Sample 3.2.1: Фрагмент аннотированных данных из набора данных OPP-115

Third Party Sharing/Collection

{'Third Party Entity': 'endIndexInSegment': 256, 'startIndexInSegment': 0, 'selectedText': 'We reserve the right to share information in order to investigate, prevent, or take action

regarding illegal activities, suspected fraud, violations of the Geekdo Terms of Service, situations involving potential threats to the physical safety of any person’, ’value’: ’Unnamed third party’, ...}

https://boardgamegeek.com/privacy

Для категории "Third Party Sharing/Collection"можно увидеть начальный и конечный индекс для предложения, извлеченного из всей политики. Благодаря структурированному формату, подробным аннотациям и меткам, которые можно использовать в обучении, было решено использовать набор данных OPP-115.

3.2.2 ACL Coling Dataset

Набор данных ACL Coling [53] был выбран для использования в основном в целях оценки. Он содержит различный набор из 1 010 юридических документов, собранных в 2014 году в формате xml. Сам корпус был создан для конференции по вычислительной лингвистике и состоит из различных типов текстов. Он является идеальной коллекцией политик для проверки прочности моделей, обученных на наборе данных OPP-115, который является более структурированным, обобщенным и устойчивым к невиданным данным. Пример того, как выглядят данные, показан ниже:

Sample 3.2.2: Пример структуры данных из набора данных ACL-Coling

<POLICY> modification_date="March 01, 2013"policy_url="http://www.zendesk.com/company/privacy"w
<SECTION>
<SUBTITLE />
<SUBTEXT>Privacy Policy

Effective as of March 1, 2013. For the prior version of our Privacy Policy, click here.

At Zendesk, we respect and protect the privacy of visitors to our website, www.zendesk.com (together with the other websites we own and control, the “Zendesk Websites”), and our customers who use our on-demand customer service support platform, tools and services offered on the Zendesk Websites (together with the Zendesk Websites, the “Service”). This Privacy Policy (“Policy”) explains how we collect and use visitors’ and customers’ information, particularly personal information, as part of the Service. The information Zendesk collects and uses is limited to the purpose for which customers engage Zendesk and other purposes expressly described in this Policy. Any discussion of your use of the Service in this Policy is meant to include your visits and other interactions with the Zendesk Websites, whether or not you are a user of Zendesk’s on-demand customer service support platform.

</SUBTEXT>
</SECTION>
<SECTION>
<SUBTITLE>Privacy Certifications </SUBTITLE>
<SUBTEXT> </SUBTEXT>
</SECTION>
</POLICY>

На рисунке 3.2.2 текст, выделенный розовым цветом, является корнем

данной политики, внутри него находятся разделы, выделенные оранжевым цветом. У каждого раздела есть подзаголовок, выделенный зеленым, и подтекст, выделенный фиолетовым.

Атрибут	OPP-115	ACL Coling
Фокус	Анализ политики конфиденциальности	Исследования в области вычислительной лингвистики
Тип контента	Политики конфиденциальности веб-сайтов	Политики конфиденциальности веб-сайтов
Аннотации	Практика использования данных (например, сбор, обмен)	Нет
Случай использования	Обучение моделей для анализа политики конфиденциальности	Широкие исследования в области НЛП и лингвистики
Формат данных	Аннотированный текст	Обычный текст, PDF
Доступность	Ограниченный доступ	Открытый доступ
Размер	115 политик	1,010
Год выпуска	2016	2013 & 2014
Поддерживается	Проект Usable Privacy Policy	Ассоциация вычислительной лингвистики

3.3 Обработка данных

Для любого анализа текста эффективность предварительной обработки данных имеет большое значение, что сказывается на производительности моделей. Особенно это важно для юридических текстов, поскольку точность представления текста играет важную роль. В этом разделе рассматриваются методы предварительной обработки, реализованные на наборах данных OPP-115 и ACL Coling, чтобы сделать их пригодными для продолжения экспериментов.

Набор данных OPP-115 потребовал обширной очистки текста, особенно потому, что для получения политики с эквивалентной меткой нам нужно было просмотреть аннотированную версию и соединить предложения вместе, чтобы получить всю политику конфиденциальности. Предварительная обработка включала в себя следующие шаги:

- **Очистка текста:** По сути, это удаление остатков HTML-тегов и неалфавитных символов. Здесь мы хотим избавиться от элементов, которые потенциально могут повлиять на анализ текста.
- **Токенизация:** Разбиение текста на лексемы. Для естественного языка этот шаг помогает разобрать текст в форме, которую модель может лучше обработать.
- **Лемматизация:** Сокращение слов до их корневой формы. Лемматизация, как правило, сохраняет контекст, а это очень важно для поддержания семантической целостности политик.

Набор данных ACL Coling требует несколько иного подхода к предварительной обработке. Учитывая природу структуры XML, процесс начинается

с разбора XML-документов, чтобы правильно извлечь текстовое содержимое. Поэтому был написан код для определения и чтения определенных элементов, выделенных в иерархии XML с помощью 3.2.2. Это подтексты разделов, поскольку именно в них хранится соответствующий юридический текст.

После того как текст извлечен, он проходит несколько дополнительных этапов очистки:

- **Normalization:** Весь текст преобразуется в строчные буквы, так как это помогает уменьшить сложность остальной обработки.
- **Обработка пробельных символов и новой строки:** Если в тексте есть лишние пробельные символы или символы новой строки, они удаляются. Это предотвращает любые ошибки при разборе, которые могут возникнуть на этапе токенизации.
- **Токенизация:** Затем текст разбивается на лексемы, чтобы преобразовать его в форму, более удобную для обработки моделью.
- **Лемматизация:** Каждая лексема приводится к своей базовой форме, сохраняя при этом семантическое значение текста.
- **Удаление пунктуации:** Пунктуация удаляется, чтобы сосредоточиться на самих словах. В противном случае они могут внести шум в модели НЛП.

Эти шаги по предварительной обработке выполняются для того, чтобы максимально повысить эффективность НЛП-моделей, используемых в данном эксперименте. Это достигается путем предоставления чистых, последовательных и содержательных текстовых данных.

3.4 Mapping GDPR Principles

Сопоставление принципов GDPR с категориями набора данных OPP-115 является следующим шагом в методологии. Поскольку мы хотим классифицировать политики как соответствующие или не соответствующие требованиям GDPR на основании того, следуют ли они принципам GDPR. В этом разделе описывается, как и почему каждая категория из набора данных OPP-115 соотносится с соответствующими принципами GDPR из статьи 5.

Категории набора данных OPP-115, перечисленные в 3.2, сопоставлены с принципами GDPR, как показано в [54]. Это сопоставление выглядит следующим образом:

- **Сбор/использование первой стороной:** Эта категория сопоставлена с законностью, справедливостью, прозрачностью, ограничением цели и минимизацией данных.
- **Обмен/Сбор третьей стороной:** Аналогично, это подразумевает законность, справедливость, прозрачность, ограничение цели и минимизацию данных.
- **Выбор пользователя/контроль:** Сопоставлено с законностью, справедливостью, прозрачностью.
- **Доступ пользователей, редактирование и удаление:** Связано с законностью, справедливостью, прозрачностью и точностью.
- **Хранение данных:** Соответствует ограничению хранения.
- **Безопасность данных:** Соотносится с целостностью и конфиденциальностью.

- **Изменение политики:** Также отображается на Законность, Справедливость, Прозрачность.
- **Международная и специфическая аудитория:** Соотносится с Законность, Справедливость, Прозрачность.
- **Не отслеживать** и **Прочее:** Эти категории не имеют прямого соответствия конкретным принципам GDPR.

Этот процесс сопоставления не только помогает обучать модели NLP, но и структурировать проверки соответствия. Мы также можем увидеть это сопоставление визуально, как показано ниже:

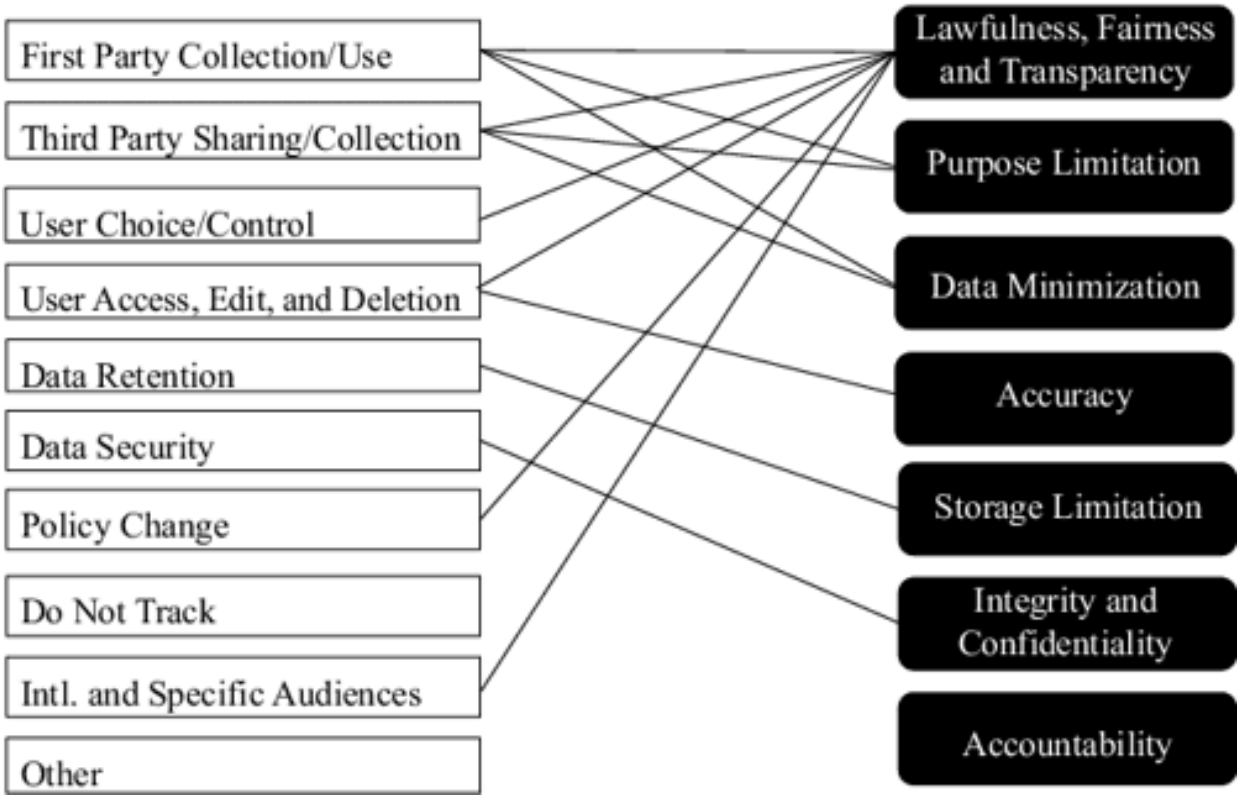

Fig. 3.1. Сопоставление категорий OPP-115 с принципами статьи 5 GDPR

3.5 Гранулярность анализа

В контексте применения НЛП для обеспечения соответствия GDPR гранулярность анализа имеет большое значение, поскольку мы применяем различные модели. Именно поэтому мы используем и сравниваем два уровня детализации: уровень предложений и уровень всей политики. Каждый уровень имеет свои преимущества и проблемы, о которых мы сейчас расскажем.

3.5.1 Анализ на уровне предложения

Преимущества:

- **Точность обнаружения:** Выявление проблем, связанных с соответствием, на уровне предложений несколько точнее, чем на уровне документов или слов. Это позволяет модели сконцентрироваться на конкретных фразах и положениях в тексте, которые относятся к принципам GDPR, поэтому она может быть очень полезна для целей проверки соответствия.

- **Контекстная релевантность:** Перенос анализа на уровень предложения позволяет учитывать окружающий контекст и убедиться, что, как бы ни работала связь Ансельмиана, она работает на уровне каждого предложения.

Трудности:

- **Потеря более широкого контекста:** Рассматривая небольшие фрагменты текста, читатель получает очень конкретные и эффективные результаты; однако стратегия не учитывает общую идею политики, что

может привести к упущению определенных закономерностей или целей, которые прослеживаются в нескольких предложениях или разделах.

- **Фрагментация:** Такая детализация может вызвать проблемы в виде фрагментации, когда анализ проводится только на основе предложений, они могут иметь ограниченное представление об общей контекстуальной схеме политики.

3.5.2 Анализ уровня политики

Выгоды:

- **Холистическое понимание:** Оценка политики как единого текста позволяет рассмотреть всю идею документа и соответствие политики требованиям GDPR. Особенно полезным является набросок общего соответствия правилам защиты данных.
- **Эффективность:** Однако в больших массивах данных удобнее оценивать целые документы, пусть и в меньшей степени, чем в более детальном варианте анализа результатов работы на уровне предложений.

Проблемы:

- **Потенциал чрезмерного обобщения:** Такой подход к анализу может привести к тому, что вы не заметите отдельные несоответствующие слова, фразы или даже абзацы во всем документе, хотя на первый взгляд он кажется абсолютно соответствующим требованиям.

- **Сложность работы с большими текстами:** Полные политики, особенно длинные и сложные, очень сложно сразу препарировать с вычислительной точки зрения, а также из-за точности и стабильности модели при работе с большими текстами.

Таким образом, предполагается использовать преимущества обоих вариантов и впоследствии выбрать тот уровень, который дает наилучший результат. Это повысит эффективность процесса проверки и даст более надежные результаты при проверке на соответствие.

3.6 Выбор модели

Выбор подходящих моделей NLP - самый важный шаг для успешного решения любой задачи машинного обучения. Особенно в области анализа юридических текстов, где важны точность и глубинные тона интерпретации. Мы рассмотрим три различные модели, и теперь внимательно изучим каждую из них, чтобы понять, почему они были выбраны.

3.6.1 SBERT

SBERT показывает хорошие результаты при получении вкраплений на уровне предложений, которые также являются более семантически обоснованными, чем обычный BERT. Это особенно полезно для некоторых упражнений, таких как сходство предложений и кластеризация; следовательно, это очень важно для классификации различий в соответствии GDPR в различных политиках конфиденциальности.

3.6.2 BERT

Одна из самых сильных сторон BERT заключается в том, что он умеет улавливать языковой контекст и структурные отношения, поэтому он очень хорошо справляется с заданиями на распознавание отношений. Это делает BERT пригодным для оценки и извлечения точной, связанной с законом информации из сложных записей, таких как политики конфиденциальности.

3.6.3 GPT

Стоит отметить, что GPT создан в основном для создания грамматически правильного и контекстуально подходящего текста. Это полезно для деятельности, предполагающей создание контента, например политик конфиденциальности или пояснительных текстов на основе GDPR, которые улучшают взаимодействие пользователей с инструментами обеспечения соответствия нормативным требованиям благодаря автоматизированным и понятным ответам.

3.7 Многозначная классификация

В данной диссертации метод классификации по нескольким меткам был использован для определения степени несоответствия каждого фрагмента текста, отдельных и многочисленных отправлений или всей ПОЛИТИКИ семи принципам, изложенным в статье 5 GDPR.

Тем не менее, в рамках GDPR каждое отдельное высказывание или даже вся политика оценивается не только с точки зрения соответствия или несоответствия, но и с точки зрения того, сколько принципов GDPR одновременно

нарушено в ней. Это реальная ситуация, когда одна политика конфиденциальности, казалось бы, учитывает одни принципы GDPR и игнорирует другие. Таким образом, каждая единица текста связана с числовым вектором длины семь, который фиксирует признание семи принципов GDPR. Каждый компонент вектора является бинарным и равен либо 0, либо 1; маленькое число 1 означает соблюдение конкретного принципа GDPR, а число 0 - нарушение.

Метки Метки для этой задачи классификации получены из принципов GDPR [55], а именно:

- Законность, справедливость и прозрачность
- Ограничение целей
- Минимизация данных
- Точность
- Ограничение хранения
- Целостность и конфиденциальность
- Подотчетность

Каждый из этих принципов формирует измерение в векторе меток для каждой точки данных. Предложение или политика аннотируется "1"для принципа, если оно соответствует требованиям GDPR для этого принципа, и "0"в противном случае.

3.8 Compliance Reporting

Эффективное соблюдение GDPR требует не только выявления потенциальных несоответствий, но и четкого информирования о них. В этом разделе описывается процесс, использованный в данной диссертации для преобразования результатов модели в комплексные отчеты о соответствии.

Процесс создания отчета о соответствии выглядит следующим образом:

- **Интерпретация результатов модели:** Все модели NLP, описанные в данной диссертации, предоставляют вероятности для всех принципов GDPR относительно того, соответствует ли данный текст, который может быть предложением или политикой, этим принципам. Таким образом, эти вероятности преобразуются в двоичные состояния, либо соответствующие, либо несоответствующие, в зависимости от определенного предела вероятности. Например, если вероятность того, что текст соответствует первому принципу "минимизации данных выше нуля. 8, он классифицируется как соответствующий. Этот порог был оптимизирован, как уже говорилось в предыдущем разделе, посвященном обсуждению кривой "точность-вызов".

- **Агрегация результатов:** После классификации отдельных предложений или разделов последним шагом является объединение результатов классификации отдельных предложений или разделов политики для получения общего отчета о соответствии уровню политики. Он включает в себя сводку всех результатов проверки принципов GDPR с представлением общей картины соответствия. Если обнаруживается нарушение какого-либо из принципов, то весь документ помечается

с указанием областей, которые следует рассматривать как несоответствующие принципам.

- **Формат отчета:** Итоговый отчет о соответствии структурирован таким образом, чтобы предоставить четкие и действенные выводы. Каждый отчет включает:
  - **Сводный раздел:** Обзор, отражающий общий статус соответствия.
  - **Подробные выводы:** Примеры выдержек из текста, подчеркивающие области, отвечающие требованиям.
  - **Рекомендации:** На основе выводов о несоответствии даются рекомендации о том, какие принципы необходимо добавить, чтобы соответствовать стандартам GDPR.

Sample 3.8.1: Пример отчета о соответствии

Отчет о соответствии GDPR для политики XYZ

Резюме: Несоответствие 2 из 7 оцениваемых принципов.

Подробные выводы:
1. Минимизация данных: Соответствует требованиям.
2. Целостность и конфиденциальность: Не соответствует требованиям.
Пример: "Данные пользователя могут храниться неограниченное время для аналитики".

Рекомендации:
- Пересмотреть политику хранения данных, чтобы привести ее в соответствие с

принципом "ограничения хранения".

Таким образом, представленная система отчетности о соответствии гарантирует, что результаты моделей будут не только понятны, но и практичны, что делает их значимыми для организаций, которым необходимо надлежащее соответствие GDPR.

3.9 Performance Metrics

Одним из наиболее важных вопросов, которые необходимо решить для обеспечения корректной работы и надежности NLP-модели, и который считается особенно важным в задачах соответствия GDPR - поскольку этот вопрос является достаточно критичным - является корректная оценка последней. В этом разделе подробно описаны метрики, используемые для оценки производительности моделей, применяемых в данной диссертации: Как уже говорилось, к метрикам производительности относятся: потери, точность, прецизионность, отзыв и F1-score. Все они выбраны в соответствии с мнением, что данные показатели могут дать лучшее понимание областей применения различных моделей и их потенциала.

Модели оцениваются с помощью следующих показателей:Модели оцениваются с помощью следующих показателей:

- **Loss:** Во время обучения важно отслеживать функцию потерь для лучшего понимания процесса обучения и сходимости модели. Если говорить о специфике данной диссертации, то общие метрики потерь важны

при настройке гиперпараметров архитектуры модели. При этом снижение потерь прямо пропорционально улучшению восприятия моделью специфики GDPR.

- **Точность:** После подготовки наборов данных точность полезна для измерения общей производительности модели в процессе обучения. В каждом из них точность полезна для оценки роста производительности с увеличением количества эпох и для сравнения простых моделей с улучшенными версиями, включенными позже.

- **Точность:** В юридической сфере оставление текста, отвечающего требованиям, в категории несоответствующих влечет за собой издержки, известные как ложные срабатывания, за которые организации сталкиваются с проверками и аудитами. Поэтому в процессе обучения модели особое внимание уделялось точности. А такие изменения, как взвешивание классов и введение регуляризации, позволили повысить точность модели при сохранении других характеристик.

- **Recall:** Показатель Recall должен быть очень высоким, чтобы охватить все возможные несоответствия. В моих экспериментах запоминание представляет интерес, особенно при измерении производительности модели на неизвестных тестовых наборах. Важно гарантировать, что модели достаточно точно определяют как можно больше случаев, чтобы не пропустить множество действительно положительных случаев.

- **F1-Score:** Из-за дисбаланса классов после таких этапов предварительной обработки, как удаление стоп-слов и стебли, использование F1-score было неизбежным. Он дает наглядное сравнение точности и

запоминания каждой модели. Это помогло определить, какая модель является лучшей из всех моделей, не перегружая при этом класс большинства.

В совокупности все эти показатели гарантируют, что модели, разработанные для проверки соответствия GDPR, прошли тестирование и валидацию, что в свою очередь обеспечивает уверенность в использовании моделей для автоматизированной проверки соответствия.

Глава 4

Реализация

В этой главе описываются детали реализации моделей машинного обучения, используемых в данной диссертации, включая SBERT, BERT и GPT. Каждая модель была обучена как на уровне предложений, так и на уровне всей политики с использованием набора данных OPP-115 с метками. Затем модели были сохранены и протестированы на немаркированных политиках из набора данных ACL Coling, чтобы оценить их обобщенность и производительность. Кроме того, в ходе реализации были проведены эксперименты с вкраплениями GPT-3 с последующим обучением этих вкраплений на классификаторе.

4.1 Подготовка данных

Сначала предложения и сами полисы были аннотированы бинарными векторами, которые могут описывать соответствие семи принципам GDPR, что позволило обучить отдельные модели для анализа на уровне предложений и полисов. Все данные проходили через токенизацию и кодирование,

чтобы соответствовать потребностям конкретной модели; например, SBERT, BERT и GPT требовали разных форматов данных для работы.

Присвоение и распределение меток Примечательно, что каждое предложение или политика были описаны с помощью двоичного 7-вектора, где каждый компонент равен 1 в позиции, относящейся к соответствующему принципу GDPR. Каждый из них приводится к вектору; если значение элемента равно 1, то организация придерживается соответствующего принципа, а если равно 0, то организация не придерживается этого принципа. Распределение этих меток по набору данных следующее:Распределение этих меток по набору данных следующее:

- **GDPR Принцип 1 (законность, справедливость и прозрачность):** 8460 экземпляров предложений и 115 экземпляров политики отмечены как соответствующие требованиям.
- **Принцип 2 GDPR (Ограничение цели):** 6209 примеров предложений и 115 примеров политики, отмеченных как соответствующие требованиям.
- **GDPR Принцип 3 (Минимизация данных):** 6209 предложений и 115 примеров политики отмечены как соответствующие требованиям.
- **GDPR Принцип 4 (Точность):** 646 примеров предложений и 90 примеров политики, отмеченных как соответствующие требованиям.
- **GDPR Принцип 5 (Ограничение хранения):** 396 экземпляров предложений 76 экземпляров политики отмечены как соответствующие требованиям.

- **GDPR Принцип 6 (Целостность и конфиденциальность):** 1000 экземпляров предложений 102 экземпляра политики, отмеченных как соответствующие требованиям.

- **GDPR Принцип 7 (Подотчетность):** 0 предложений и примеров политики, отмеченных как соответствующие требованиям.

Такое распределение меток показывает, что принципы разделены на подтемы, представляющие интерес в данных, и что обучение моделей может быть проблематичным, особенно в принципах с относительно небольшим количеством положительных примеров. Таким образом, разница в частоте меток подчеркивает тот факт, что модели должны быть способны обнаруживать нечастые, но важные проблемы соответствия, включая Точность и Ограничение хранения.

Перед обучением набор данных OPP-115 был подготовлен с двумя различными гранулярностями:

1. **уровень предложения:** В случае с политиками конфиденциальности каждое предложение аннотировалось с точки зрения того, насколько оно соответствует представленным принципам GDPR, что и послужило обучающими данными для моделей на уровне предложений.

2. **Уровень политики:** Полные политики конфиденциальности были классифицированы на основе соответствия GDPR, четные или нечетные, как единое целое для обучения моделей уровня политики.

4.2 Обучение и сохранение моделей

В процессе обучения моделей необходимо было создать структуры для нейронных сетей, так как вкрапления документов предварительно обрабатывались с помощью моделей SBERT, BERT и GPT. Все перечисленные модели были обучены на наборе данных OPP-115, разделенном на тренировочный и валидационный наборы. После обучения все модели были сериализованы стандартным образом.

4.2.1 SBERT

Sentence Level

Таким образом, версия модели SBERT all-MiniLM-L6-v2, которая была принята для обеспечения семантически совпадающих вкраплений предложений. Эта возможность важна для обеспечения соответствия GDPR на уровне предложений.

Сначала с помощью SBERT были сгенерированы вкрапления из набора данных OPP-115, которые затем были взяты классификатором в качестве признаков. Эти вкрапления также сохраняют всю информацию о предложении и его пригодности для высокоэффективной классификации.

Далее, для управления этими вкраплениями, а также связанными с ними метками соответствия GDPR, которые представляют собой 7-элементный двоичный вектор для каждого предложения, был создан новый пользовательский набор данных PyTorch, названный SBERTEmbeddingDataset. С помощью этого набора данных во время обучения загрузка и сортировка данных становится проще.

Листинг 4.1: Скод фрагмента пользовательского набора данных

```
sentence_sbert_model = SentenceTransformer('all-MiniLM-L6-v2')

def sentence_sbert_generate_embeddings(texts):
    embeddings = sentence_sbert_model.encode(texts, convert_to_tensor=True)
    return embeddings

class SBERTEmbeddingDataset(Dataset):
    def __init__(self, texts, labels):
        self.embeddings = sentence_sbert_generate_embeddings(texts)
        self.labels = labels

    def __len__(self):
        return len(self.embeddings)

    def __getitem__(self, idx):
        return {
            'embeddings': self.embeddings[idx],
            'labels': torch.tensor(self.labels[idx], dtype=torch.float)
        }
```

Что касается обучения, то оно используется для создания нейросетевого классификатора, включающего LSTM и слои внимания, которые позволяют использовать вкрапления предложений, предназначенные для многопометной классификации. Процесс обучения контролировался циклом, в котором каждая итерация, называемая эпохой, подразумевала обучение на обучающем множестве, а затем на валидационном множестве.

Для оценки эффективности работы модели использовались следующие метрики. К ним относятся точность, которая представляет собой отношение числа правильных предсказаний к числу образцов, а также точность, отзыв и F1 score, которые, особенно в доменах с несколькими метками, где соответ-

ствие каждому принципу рассматривается как уникальная метка, являются более точными показателями эффективности модели.

Впоследствии модель была сохранена с помощью механизма сохранения PyTorch, где хранятся параметры модели с целью ее тиражирования и проведения дополнительных тестов в зависимости от модели, либо создания производственной модели.

Такая детальная реализация модели SBERT на уровне отдельных предложений гарантирует, что каждое предложение будет распознано как соответствующее GDPR с помощью достаточно эффективных методов NLP для анализа юридических текстов.

Уровень политики

На уровне политики та же модель Sentence-BERT, то есть модификация all-MiniLM-L6-v2, была использована для обработки необработанных текстов политик конфиденциальности. Этот подход был задуман для проверки соответствия GDPR полных документов и, следовательно, позволял получить макроскопическое представление об этих требованиях.

В случае с политиками SBERT использовался для получения вкраплений полных политик. Это возможность обработки больших блоков текста для получения общей "сути"или семантического содержания документа; это необходимо для оценки общего соответствия каждой политики.

Для анализа использовался один и тот же пользовательский набор данных SBERTPolicyEmbeddingDataset, состоящий из меток соответствия GDPR каждой политики и бинарных векторов.

Обучение модели проводилось с помощью классификатора с комбина-

цией слоев LSTM и внимания, что повышает совместимость с результатами работы SBERT. Такая настройка позволяет дополнительно обогатить знания о тонкостях соответствия на уровне политик в зависимости от контекста и тематики.

Для оценки производительности обученной модели в реальном мире она была протестирована на разделенном тестовом наборе. Основные показатели эффективности, которые применялись при оценке модели, включают точность, прецизионность, отзыв и F1-score.

После проверки модель была сохранена для повторного использования или для переноса в другие среды, особенно в операционную среду.

Эта реализация также доказала, что SBERT предназначен не только для одного предложения, но и для всего документа, что становится полезным при поиске соответствия GDPR во всем содержании политик конфиденциальности. Это особенно важно при оценке соответствия, поскольку подразумевает комплексный подход при решении задач организации, чтобы ее политики соответствовали всем необходимым требованиям законодательства.

4.2.2 BERT

Уровень предложения

В результате сложности соблюдения GDPR на уровне предложений для решения этой задачи была использована модель BERT (двунаправленные кодирующие представления из трансформаторов) с bert-base-uncased расположением. Благодаря мощным способностям BERT к пониманию языка, он очень полезен при анализе сложных юридических текстов.

Чтобы повысить эффективность обучения, исходные данные были под-

готовлены путем токенизации с помощью токенизатора BERT. Все предложения из набора данных OPP-115 были предварительно обработаны для ввода в BERT, для чего потребовалось следующее: текст был разбит на лексемы максимальной длины 512, а в тех случаях, когда строки были короче, добавлялась прокладка, а если строки были длиннее, они усекались.

Разбивка данных производилась в соотношении $80:20$, при этом первые 80% использовались для обучения, а последние 20% - для оценки.

В частности, для обучения использовался пользовательский WeightedTrainer, который позволяет настраивать функцию потерь на основе весов классов, что было направлено на сбалансированное отношение ко всем классам, включая менее частые.

При оценке модели использовались различные метрики, такие как accuracy, precision, recall, F1-score. Каждая из этих метрик позволяет оценить работу модели в классификации уровня соответствия GDPR.

Модель и токенизатор после обучения были сохранены для воспроизведения, а также для дальнейшего использования модели для тестирования или развертывания. Результаты обучения и оценки также были сохранены для учета и демонстрации эффективности модели.

Такая широкая настройка демонстрирует тщательное использование BERT для текстового анализа на уровне предложений, что означает, что каждый аспект соответствия GDPR тщательно изучен, а результаты, полученные моделью, подробно записаны для использования в дальнейших справочных целях или в реальной работе.

Листинг 4.2: Фрагмент кода набора данных политики конфиденциальности для уровня предложений

```
20 class PrivacyPolicySentenceDataset(Dataset):
21     def __init__(self, texts, labels, tokenizer, max_len):
22         self.tokenizer = tokenizer
23         self.texts = texts
24         self.labels = labels
25         self.max_len = max_len
26 
27     def __len__(self):
28         return len(self.texts)
29 
30     def __getitem__(self, item):
31         text = str(self.texts[item])
32         label = self.labels[item]
33         encoding = self.tokenizer.encode_plus(
34             text,
35             add_special_tokens=True,
36             max_length=self.max_len,
37             return_token_type_ids=False,
38             padding='max_length',
39             return_attention_mask=True,
40             return_tensors='pt',
41             truncation=True
42         )
43 
44         return {
45             'input_ids': encoding['input_ids'].flatten(),
46             'attention_mask': encoding['attention_mask'].flatten(),
47             'labels': torch.tensor(label, dtype=torch.float)
48         }
```

Всего уровня политики

Что касается уровня всей политики, то модель BERT использовалась с версией bert-base-uncased, в которой в нее подавались целые политики кон-

фиденциальности. Этот уровень анализа очень полезен для оценки общего соответствия политики GDPR, поскольку можно выявить аспекты, которые не соответствуют GDPR, если проанализировать только одно или несколько предложений.

Политики конфиденциальности обрабатывались с помощью токенизатора BERT с ограничением максимальной длины для управления размером входных текстов, обеспечивая их соответствие возможностям модели при сохранении важной информации:

Разработанный набор данных был снова разделен на обучающий и оценочный, с процентным соотношением $80:20$. Такое разделение было сделано для того, чтобы пройти тесты на невидимых данных, так как это могло бы устранить некоторые расхождения в оценке.

Используемая BERT-модель была предварительно обучена предсказывать соответствие каждой политики одной из семи категорий соответствия, что связано с принципами GDPR. Обучение включало в себя задание параметров модели для точности, прецизионности, запоминания и F1-score, которые являются специфическими и жизненно важными, когда речь идет об оценке соответствия высоким требованиям.

Эффективность предложенной модели была проанализирована по заданным контрольным точкам, а функция расчета метрик была ориентирована на производительность модели с точки зрения точности классификации политик по отношению к уровню соответствия.

После этого модель вместе с ее токенизатором была сохранена для использования в других операционных средах или для других оценок. Этот шаг обеспечивает масштабируемость модели и ее готовность к проверке соответствия GDPR в режиме реального времени.

Этот раздел служит для демонстрации всего спектра, в котором BERT применяется для полного анализа политики, что характеризует масштабы обработки и анализа, которые используются для обеспечения соответствия стандартам GDPR.

Листинг 4.3: Скод фрагмента набора данных политики конфиденциальности

```
52 class PrivacyPolicyDataset(Dataset):
53     def __init__(self, policies_text, labels, tokenizer, max_len):
54         """
55         Args:
56             policies_text (list of str): Texts of multiple policies.
57             labels (list of list of int): Labels for each policy.
58             tokenizer: Tokenizer to be used for encoding the text.
59             max_len (int): Maximum length of the tokens.
60         """
61         self.tokenizer = tokenizer
62         self.policies_text = policies_text
63         self.labels = labels
64         self.max_len = max_len
65 
66     def __len__(self):
67         return len(self.policies_text)
68 
69     def __getitem__(self, idx):
70         encoding = self.tokenizer.encode_plus(
71             self.policies_text[idx],
72             add_special_tokens=True,
73             max_length=self.max_len,
74             return_token_type_ids=False,
75             padding='max_length',
76             truncation=True,
77             return_attention_mask=True,
78             return_tensors='pt'
79         )
```

```
80
81          return {
82              'input_ids': encoding['input_ids'].flatten(),
83              'attention_mask': encoding['attention_mask'].flatten(),
84              'labels': torch.tensor(self.labels[idx], dtype=torch.float)
85          }
```

4.2.3 GPT

Sentence Level

На уровне предложений использовалась модель GPT-2, которая представляет собой генеративную предварительно обученную модель трансформации, благодаря ее хорошей производительности в генерации подробных контекстно-зависимых вкраплений. Такой анализ актуален, когда необходимо препарировать политику конфиденциальности и проверять каждое предложение на соответствие принципам GDPR.

Первая операция включала предварительную обработку данных, аналогичную токенизатору GPT-2. Этот токенизатор подготавливает текст для ввода в GPT-2, разбивая его на предложения и затем токенизируя каждое из них, чтобы включить в него прокладку нужного размера.

Таким образом, модель GPT-2 была использована для классификации заданных предложений на соответствующие или несоответствующие GDPR, с соответствующей специфической схемой обучения. Это потребовало изменения функции потерь для решения проблемы дисбаланса классов с помощью взвешенных потерь.

Класс Hugging Face Trainer был использован для создания режима обу-

чения для моделей NER. Настройка Hugging Face Trainer для достижения взвешенных потерь по классам означает, что функция потерь была адаптирована для лучшей оценки модели на основе всех классов, независимо от количества точек данных для каждого класса.

Точность, точность, отзыв и F1 score были рассчитаны после обучения модели для полной оценки. Эти метрики показывают количество истинно положительных, истинно отрицательных, ложно положительных и ложно отрицательных результатов, чтобы определить, насколько точно данная модель различает сайты, соответствующие GDPR, и сайты, не отвечающие требованиям.

После проверки обученная модель и токенизатор сохраняются, чтобы их можно было использовать или применять для других реальных проверок на соответствие GDPR. Кроме того, результаты обучения и журналы были заархивированы для целей документации и анализа:Кроме того, результаты обучения и журналы были заархивированы для целей документации и анализа:

Таким образом, такой комплексный подход к реализации GPT-2 на уровне предложения говорит о возможности глубокого анализа текста данной моделью на предмет соответствия GDPR с применением передовых методов НЛП для приведения политик конфиденциальности в соответствие с нормами.

Уровень политики конфиденциальности

Более конкретно, GPT-2, продемонстрировавший сильные возможности в рассуждениях и создании текста, напоминающего созданный человеком

контент, был использован на общем уровне политик для анализа соответствия политик конфиденциальности GDPR. В данном подходе GPT-2 используется специально для анализа контекста и оценки больших текстовых отрывков, что подходит для данного типа анализа.

Набор данных, сформированный из всех политик конфиденциальности, был получен с помощью токенизатора GPT-2, который гарантирует, что каждая политика будет точно токенизирована и закодирована. Что касается токенизации, то особое внимание было уделено ее эффективной реализации, поскольку возможности были ограничены количеством слов.

Обучение проводилось на версии модели GPT-2, приспособленной для классификации последовательностей по нескольким меткам. Указанная модель была разработана для точного прогнозирования семи принципов GDPR с использованием сигмоидальной функции для вероятностно-независимых логитов модели.

Для уменьшения дисбаланса классов в процессе обучения проводилось взвешенное обучение. Для оценки эффективности моделей использовались различные показатели, такие как точность, прецизионность, отзыв, F1-score.

Эффективность предложенной модели также проверялась с помощью набора измеримых критериев, позволяющих оценить эффективность модели с точки зрения ее точности в классификации соответствия GDPR и ее улучшения.

После проверки обученная модель и токенизатор были сохранены для непосредственного использования или для реального оперативного тестирования на соответствие GDPR.

Проработанная схема гарантирует, что GPT-2 хорошо подходит для применения на полных фрагментах текстовых документов с политиками конфи-

денциальности, благодаря своей эффективной способности к контекстному пониманию, обеспечивая целостный и точный аудит соответствия политикам GDPR.

4.3 Тестирование на неразмеченных данных

Сохраненные модели для обоих уровней детализации помогают применить решения к набору данных ACL Coling, который содержит немаркированные политики. Этот шаг был важен для определения предсказательной способности моделей и того, насколько хорошо они будут работать в реальных ситуациях.

4.3.1 SBERT

Уровень предложения

На последнем этапе реализации модель SBERT используется для оценки отдельных предложений политики конфиденциальности на соответствие конкретным принципам GDPR. Чтобы описать этот подход, необходимо выполнить несколько операций, таких как загрузка весов модели, генерация предсказаний на основе предложений и общая оценка политики.

Для загрузки предварительно обученного классификатора SBERT используется пользовательская функция, позволяющая убедиться, что модель находится в режиме оценки, что очень важно при составлении прогнозов без изменения изученных параметров модели.

Функция sentence_sbert_classify_policy берет политику, разбивает предложения в политике, извлекает вкрапления и использует классификатор,

чтобы перейти к проверке соответствия GDPR. Функция присваивает каждому предложению метки GDPR и собирает как сами метки, так и вероятности того, что они верны.

Результаты, полученные от классификатора, затем суммируются, чтобы предложить более комплексную оценку соответствия текущей политики. Такие меры заключаются в представлении результатов классификации, относящихся к каждому предложению с присвоенными соответствующими метками GDPR, и таких специфических меток GDPR, выявленных во всем документе, чтобы предоставить некоторую практическую информацию о соответствии политики GDPR.

Это позволяет не только провести анализ на уровне предложений, но и скомпилировать полученные результаты для оценки общего соответствия Политики. Таким образом, с помощью использования модели SBERT в предложенном виде диссертация показывает применимость высокоразвитых методов в области НЛП в реальной среде оценки соответствия GDPR.

Всего уровня политики

На уровне всей политики обученная модель SBERT используется для определения соответствия всей политики конфиденциальности GDPR. При этом оценивается документ с несколькими предложениями, что позволяет использовать возможности SBERT для анализа больших текстов.

К изображениям применяется несколько методов предварительной обработки, однако наиболее важным начальным шагом является загрузка предварительно обученного классификатора SBERT, чтобы он был готов к выводам. Модель готовится к выполнению требований, связанных с оценкой

целых политик, при этом следя за точностью и скоростью работы.

Функция для классификации политик - policy_sbert_classify_policy, которая выполняет анализ каждой политики для получения прогноза соответствия GDPR. Именно эта функция вычисляет вкрапления, задействует классификатор для прогнозирования соответствия, а также применяет определенный порог и выявляет соответствующие GDPR метки.

Поскольку классификация выполняется в режиме реального времени, конечный результат агрегируется и представляет собой обзор соответствия политики GDPR. Это включает в себя общие и конкретные прогнозы для всех политик, где дается диаграмма с уровнем соответствия, а также все конкретные метки GDPR, которые были определены.

В этом разделе показано, как использование SBERT для применения его к целым политикам конфиденциальности также приводит к гораздо более детальному и глубокому анализу соответствия GDPR. Применяя новые методы НЛП, диссертация демонстрирует способ использования методов машинного обучения для улучшения системы проверки соответствия законодательству.

4.3.2 BERT

Уровень предложения

Для анализа на уровне предложений используется модель BERT (Bidirectional Encoder Representations from Transformers), позволяющая классифицировать соответствие GDPR всех предложений в заданных политиках конфиденциальности. Такой детальный подход позволяет отличить конкретную проблему соответствия политике от общей, что является очень важным

аспектом при анализе политики.

Чтобы использовать возможности BERT, для построения конвейера был загружен предварительно обученный BERT вместе с его токенизатором. Модель обучается с целью предсказания скрытых состояний, которые в дальнейшем могут быть использованы для анализа или классификации высокого уровня.

Конвейер обучается с помощью функции ‘pipeline‘ в Hugging Face, причем используется только конвейер классификации. Этот конвейер упрощает процесс классификации предложений благодаря тому, что и модель, и токенизатор могут быть включены в конвейер в виде вызываемой функции.

Функция классификации вызывается как sentence_bert_classify_policy, которая запускает конвейер для каждого предложения. Он устанавливает фильтр, чтобы узнать, относится ли конкретное предложение к определенному тегу GDPR, в соответствии с оценкой вероятности модели.

Исходя из данного представления, получив данные, результаты обобщаются и представляются таким образом, чтобы было понятно, соответствует ли каждое предложенное предложение требованиям стандартного английского языка или нарушает их. Это влечет за собой прогнозы для всех предложений, а также краткое описание новых меток GDPR, встречающихся в политике.

В этой части BERT реализован и объяснен на уровне предложений для оценки соответствия политик GDPR, что еще раз доказывает эффективность интеграции политик на уровне корпуса и интерпретации результатов на основе классификации отдельных предложений.

Уровень политики

Модель BERT, которая, как сообщается, хорошо понимает языковой контекст, используется на уровне политик, чтобы дать общую оценку соответствия GDPR всех политик сверху донизу и всех политик конфиденциальности в целом. Такой подход позволяет получить широкую картину соответствия, которую не может показать анализ отдельных предложений из-за некоторых тонкостей.

Первый шаг включает в себя подготовку модели, полученной из предварительно обученной модели BERT и токенизатора. Модель настраивается таким образом, чтобы возвращать скрытые состояния: Они могут быть использованы для дальнейшего анализа или для улучшения классификационного бита.

Конвейер классификации определяется с помощью функции Hugging Face's pipeline. Этот конвейер помогает эффективно классифицировать текстовые данные с помощью модели.

Функция policy_bert_classify_policy специально предназначена для прогона классификационного конвейера через все политики. Он оценивает каждую политику, накладывает фильтр на модель для определения релевантности и указывает метки соответствия GDPR.

Результаты классификации затем суммируются и выводятся в отчет, чтобы предложить расширенную информацию о соответствии политики. Для этого составляется список идентифицированных меток GDPR и соответствующие уровни доверия. Этот процесс не только обеспечивает всестороннюю оценку на уровне политики, но и консолидирует эти результаты для получения общей картины соответствия политики. Таким образом, используя

BERT, диссертация демонстрирует, как современные инструменты NLP могут эффективно применяться для решения практических задач в реальных условиях анализа соответствия GDPR.

4.3.3 GPT

Уровень предложения

На уровне предложений GPT, особенно модель GPT-2, используется для прогнозирования соответствия GDPR каждого предложения в политике конфиденциальности. Именно поэтому данный подход, при котором мы фокусируемся на анализе каждого предложения в отдельности, лучше всего подходит в сочетании с GPT-2, так как эта модель лучше всего подходит для понимания контекста и создания более тонких интерпретаций текстов.

Первый процесс заключается в загрузке модели GPT-2 и прилагаемого к ней токенизатора. Модель была создана специально для генерации скрытых состояний, которые помогают интерпретировать типы создаваемых структур и улучшают общую производительность классификатора.

Конвейер классификации создается с помощью функции ‘pipeline‘ из Hugging Face. Эта настройка также облегчит применение модели к текстовым данным, чтобы повысить эффективность классификации.

Функция sentence_gpt_classify_policy берет каждое из предложений и пропускает его через конвейер классификации. Он оценивает каждое предложение по релевантности, используя установленные пороги доверия к модели, и итоговые записи будут иметь метки соответствия GDPR.

После классификации результаты суммируются и отображаются таким образом, чтобы можно было получить предварительную информацию о со-

ответствии каждого предложения законодательству. Это включает в себя отображение статистики результатов классификации каждого предложения и всех новых меток GDPR, которые были найдены в документе.

Этот раздел также демонстрирует, как GPT можно использовать на самом низком уровне, чтобы проанализировать конкретные аспекты соответствия GDPR в политиках конфиденциальности и показать способность генерировать ценные и конкретные рекомендации на основе анализа классификации предложений.

Уровень политики конфиденциальности

Использование GPT и, в частности, GPT-2 на уровне политики означает оценку соответствия GDPR целых разделов текстов, целых политик конфиденциальности, фактически. При этом используется способность GPT-2 к глубокому обучению для анализа текстовых структур, что позволяет Comply выявлять проблемы соответствия в целых документах.

В первую очередь необходимо загрузить предварительно обученную модель GPT-2 и токенизатор для предварительной обработки текста. Параметры модели изменяются, чтобы получить желаемые скрытые состояния, которые позволят лучше анализировать текст.

С помощью виджета Hugging Face 'pipeline' настраивается сквозной конвейер классификации, поскольку он позволяет легко применять модель. Такая настройка улучшает обработку текстовых данных, поскольку небольшие, но значительные текстовые данные обрабатываются маленькой сетью, а большие, но незначительные текстовые данные - большой сетью.

Вызывающей функцией в этом конвейере является функция класси-

фикации policy_gpt_classify_policy, которая должна обрабатывать каждую политику. Эта функция просматривает полный текст каждой политики и фильтрует предсказания, которые пересекают заданный порог, после чего возвращаются соответствующие метки соответствия GDPR.

После этого формируется и представляется сводка результатов, чтобы дать представление о статусе соответствия политики. Это варьируется от перечисления всех воспринятых меток GDPR с соответствующим уровнем доверия до глубокого понимания политики в соответствии с правилами GDPR.

В этом разделе показано, как на уровне политики GPT-2 используется для сравнения и оценки соответствия GDPR целых политик конфиденциальности, и этот пример отлично доказывает способность GPT-2 предоставлять конкретные и ценные рекомендации, основанные на тщательном анализе текста.

4.4 Single Principle Checker

Программа Single Principle Checker была разработана для решения проблемы обнаружения случаев несоответствия GDPR на уровне предложений, касающихся только отдельных принципов. Эта реализация была специфична для принципа "Ограничение хранения"в соответствии с законом GDPR, который является очень важным.

Обзор модели

Модель, использованная в данной работе, представляет собой вариант SBERT, называемый all-MiniLM-L6-v2, который позволяет создавать инди-

видуальные числовые представления предложений и, следовательно, вкраплений. Последние отражают более тонкую семантику текста, необходимую для оценки соответствия.

Подготовка данных

Вкрапления и метки управлялись пользовательским классом набора данных под названием SBERTEmbeddingDataset. В этом наборе данных каждое предложение превращалось в высокоразмерный вектор, который отражал всю семантическую информацию, содержащуюся в предложении, а также метку, которая была либо 1, если принцип ограничения хранения соблюдался, либо 0, если не соблюдался.

Процесс обучения

В процессе обучения использовался подход бинарной классификации, при котором для оценки соответствия использовалась архитектура нейронной сети с LSTM-слоями и вниманием. Такая настройка вкраплений позволила модели найти некоторые отличительные характеристики в отображениях фигур, что улучшило ее способность обнаруживать намеки на соответствие или его отсутствие.

Оптимизация и оценка

Эти гиперпараметры включают скорость обучения и эпохи, и для того, чтобы получить наилучшую конфигурацию, дающую наилучшую производительность с точки зрения F1 score, которая идеально подходит для бинарных проблем классификации, обучение было экспериментально проверено при

различных конфигурациях. Таким образом, были получены такие показатели, как точность, прецизионность, отзыв и F1-score, чтобы дать общее представление о производительности модели. Было определено оптимальное количество соседей, после чего модель была обучена и сохранена для данной конфигурации модели.

Детали эксперимента

Далее были проведены дополнительные эксперименты по отработке модели в подходящем стиле, пригодном для тщательного тестирования. Экспериментировались различные скорости обучения, включая $5 * 10^{(-5)}, 1 * 10^{(-5)}, 1 * 10^{(-4)}$, и количество эпох, включая $5, 10, 15$, а также пороги принятия решений $0.3, 0.5, 0.7$. Результаты показали, что функции модели различаются в зависимости от их выбора - это подтвердило факт настройки модели на принципы соответствия.

4.5 Эксперименты по встраиванию OpenAI

Кроме того, были получены вкрапления API GPT-3, которые были обработаны для классификации. В этом эксперименте снова учитывались фундаментальные генеративные возможности GPT-3 для дальнейшего повышения точности классификации.

В этом разделе описывается, как вкрапления OpenAI, и в частности модель Ada, могут быть использованы для классификации заданных предложений, основанных на политике конфиденциальности, на предмет их соответствия GDPR. Эксперименты включают: загрузку файла контрольных точек, настройку конвейера классификации, обучение нейронной сети и оценку

производительности.

Импорт модели Ada и процессов ее токенизатора из API OpenAI необходим, так как они используются для получения подходящих вкраплений, выражающих значения текстов политики конфиденциальности.

Для передачи текстов и получения соответствующих вкраплений из Ada-модели создана функция. Эти вкрапления служат первичными входами для нейросетевого классификатора.

Это необходимо для того, чтобы в процессе обучения различные политики имели свои собственные вкрапления и метки. Это облегчает сортировку данных во время обучения или проверки модели на последующих этапах.

Для определения соответствия или несоответствия данного элемента GDPR на основе вкраплений, созданных предварительно обученной языковой моделью, предлагается использовать многослойный перцептрон. Он включает линейные слои, которые в нейронных сетях принято обозначать весами, а также функции активации.

Здесь описана последовательность настройки обучения и проверки, а также определены необходимые компоненты, такие как вычисление потерь и такие метрики производительности, как точность, прецизионность, отзыв и F1 score.

4.6 Оценка модели

Наконец, что не менее важно, эффективность каждой модели была оценена на основе вышеупомянутых критериев: точности, отзыва, F1-score и точности. Результаты мы рассмотрим в следующей главе **??**. Они сравнивались для того, чтобы определить, насколько эффективно ascus|evaluate обу-

чение на уровне предложений по сравнению с обучением на уровне политик. Также было проведено сравнение производительности разработанной модели с вкраплениями GPT-3 и без них, чтобы определить вклад вкраплений GPT-3 в задачу классификации.

В этой главе читателю предлагается пошаговое описание стратегии реализации NLP-моделей, которая используется для эксперимента, проводимого в рамках данной исследовательской работы. Использование данной методологии гарантирует, что полученные результаты могут быть легко повторены, что позволяет убедиться в надежности применения полученных результатов для различных оценок соответствия GDPR.

Глава 5

Выводы

В данной диссертации основное внимание уделялось использованию НЛП для автоматизации проверки соответствия GDPR. Это начинание было продиктовано такими факторами, как растущая сложность юридических вопросов и необходимость эффективного применения законов о защите данных в организациях. В ходе комплексного анализа использовались самые современные модели НЛП, включая SBERT, BERT и GPT-2, на двух уровнях гранулярности: с одной стороны, на уровне предложений, а с другой - на уровне всей политики.

5.1 Обзор выводов

Поскольку целью данного исследования была оценка NLP-моделей для определения степени соответствия политик конфиденциальности GDPR, путем тщательного эксперимента и оценки было получено несколько ключевых результатов:

- **Эффективность модели:** Из оценки моделей можно сделать вывод,

что, хотя все модели были достаточно эффективны в той или иной степени, SBERT лучше всего проявила себя на уровне предложений, обеспечив высокую точность и запоминание вопросов соответствия. Модели BERT и GPT-2 также показали приемлемые результаты, кроме того, для BERT были выявлены области относительно высокой эффективности, в основном, при анализе отношений в тексте. Следующие баллы суммируют результаты работы:

  - **SBERT:**
    * **Уровень предложения:** Accuracy: 0.57, Precision: 0.54, Recall: 0.42, F1-score: 0.43
    * **Уровень политики:** Accuracy: 0.60, Precision: 0.76, Recall: 0.85, F1-score: 0.80
  - **BERT:**
    * **Уровень предложения:** Accuracy: 0.52, Precision: 0.72, Recall: 0.42, F1-score: 0.52
    * **Уровень политики:** Точность: 0.43, Precision: 0.80, Recall: 0.75, F1-score: 0.77
  - **GPT-2:**
    * **Уровень предложения:** Accuracy: 0.62, Precision: 0.74, Recall: 0.46, F1-score: 0.55
    * **Уровень политики:** Accuracy: 0.43, Precision: 0.79, Recall: 0.75, F1-score: 0.77

- **Гранулярность анализа:** На предоставленном уровне анализа предложений были получены более точные результаты по вопросам соот-

ветствия, что позволило провести более тщательное изучение каждого пункта соглашения. Тем не менее, подход на уровне всей политики позволил получить более общие модели соответствия, которые можно проконтролировать на детальном уровне. Сравнивая результаты, полученные на двух разных уровнях детализации, можно отметить, что каждый из них особенно актуален для определенного набора мер регулирования соответствия.

- **GPT-3 Embeddings:** Благодаря встраиванию GPT-3 улучшилась контекстная осведомленность моделей, что привело к улучшению результатов прогнозирования. Тем не менее, вкрапления GPT-3 продемонстрировали оптимальное использование вычислительных ресурсов для повышения точности и побитового понимания текста, а также точности и запоминания, достигнутых при анализе довольно тонких рисков несоответствия.

5.2 Discussion

В целом, применение НЛП для проверки соответствия GDPR полно как возможностей, так и рисков. Плюсы моделей НЛП, особенно использующих вкрапления GPT-3, - сохранение семантического смысла, контекстуальность и сочетание обширных предварительно обученных знаний. Эти модели могут значительно повысить скорость и эффективность проверок на соответствие нормативным требованиям, а значит, сократить число тех, кто просто просматривает данные вручную.

Однако проблемы остаются. Использование вкраплений GPT-3 требует

больших временных и вычислительных затрат, а API, связанные с использованием GPT-3, и стоимость являются другими изученными ограничениями. Однако предсказуемость и интерпретируемость - две главные проблемы таких методов, поскольку очень важно понимать, как модель пришла к такому решению и соответствует ли оно закону.

Будущая работа в этом направлении должна быть направлена на дальнейшее выявление и решение указанных проблем путем рассмотрения гибридных моделей, систем, основанных на правилах, и методов повышения объяснимости НЛП-систем. Таким образом, встраивание и модели должны постоянно обновляться в соответствии с современными стандартами, которые установлены в законодательстве и использовании языка сегодня.

5.3 Практическая значимость

Результаты исследования свидетельствуют о том, что руководители организаций, желающие улучшить свои программы по соблюдению GDPR, могут рассмотреть возможность реализации вышеуказанной меры. Внедряя и внедряя решения и продукты на основе НЛП, организации могут значительно улучшить управление соответствием нормативным требованиям, своевременно и правильно выполняя все предписания и связанные с ними процессы. Модели и методы, предложенные в данной диссертации, могут быть интегрированы в существующие системы обеспечения соответствия и, таким образом, представляют собой относительно простые в реализации и эффективные решения.

Кроме того, обсуждение показало, что существует необходимость интеграции специалистов в области права и ученых в области НЛП для развития

знаний, полученных в этих двух областях. Такое партнерство может привести к созданию более совершенных и контекстуально улучшенных инструментов соответствия, которые, в свою очередь, укрепят меры по защите данных и частной жизни.

5.4 Ответы на вопросы исследования

В данной диссертации ставится задача ответить на несколько ключевых вопросов, и полученные результаты дают некоторые ответы и понимание выбранных вопросов исследования:

- **Насколько эффективны модели НЛП, включая GPT-3, для автоматизации выявления проблем соответствия GDPR в политике конфиденциальности данных организации?** Результаты исследования показывают, что модели НЛП очень эффективны и точны, при этом основное внимание уделяется модели SI, включающей SBERT и BERT, для понимания соответствия требованиям. И без того приятное преобразование PDF в текст также улучшено с помощью вкраплений GPT-3, которые обеспечивают глубокое семантическое понимание и осознание контекста.

- **Каковы ограничения текущих технологий НЛП в интерпретации и обеспечении соответствия GDPR, и как эти ограничения можно устранить?** Основные недостатки заключаются в сложности вычислений и необходимости использования менее сложных с когнитивной точки зрения моделей, таких как GPT-3, а также в проблемах интерпретируемости. Преодолеть эти ограничения можно, увеличив скорость

вычислений, сделав модели более доступными и придумав способы, с помощью которых модели могут быть легко поняты.

- **Какую роль могут сыграть инструменты на базе NLP в поддержке специалистов по соблюдению нормативных требований и юристов в обеспечении соответствия GDPR?** Также показано, что инструменты на базе NLP могут помочь специалистам по соблюдению нормативных требований, автоматизируя такие задачи, как правильное выявление проблем с соблюдением нормативных требований в организации, что снижает объем работы специалиста и повышает уровень надежности проверок соблюдения нормативных требований. Эти инструменты могут помочь в обеспечении соответствия, что, в свою очередь, освободит время специалистов по правовым вопросам для выполнения более сложных задач.

5.5 Вклад в область

Применительно к области конфиденциальности данных и соответствия нормативным требованиям, данная диссертация помогает установить использование передовых методов НЛП в сфере решения повторяющихся, сложных интеллектуальных задач. Сравнение различных моделей дает представление об их преимуществах и недостатках, а также помогает наметить направления дальнейших исследований в этой области. Таким образом, использование вкраплений GPT-3 можно рассматривать как серьезное усовершенствование, демонстрирующее, как использование последних достижений в области обработки языка ИИ может помочь в обработке юридических тек-

стов.

В целом, данная диссертация показывает, что НЛП может привнести положительные изменения в решение вопросов, связанных с регулированием GDPR. Благодаря использованию передовых языковых моделей и методов встраивания можно лучше понять нормативную среду и справиться с ней, что повысит информационную прозрачность и дружественный подход пользователей к конфиденциальности данных.

Список использованной литературы

[1] J. Hirschberg и C. D. Manning, «Advances in natural language processing,» *Science*, т. 349, № 6245, с. 261—266, 2015.

[2] *AI Review for Data Processing Agreements (DPAs) — legalontech.com*, https://www.legalontech.com/contracts/data-processing-agreement-dpa, [Accessed 13-06-2024].

[3] O. Amaral Cejas, S. Abualhaija и L. Briand, «ML-based Compliance Verification of Data Processing Agreements against GDPR,» English, сент. 2023. url: https://orbilu.uni.lu/handle/10993/55408.

[4] C. Meehan, K. Mrini и K. Chaudhuri, «Sentence-level Privacy for Document Embeddings,» в *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, S. Muresan, P. Nakov и A. Villavicencio, ред., Dublin, Ireland: Association for Computational Linguistics, май 2022, с. 3367—3380. DOI: 10.18653/v1/2022.acl-long.238. url: https://aclanthology.org/2022.acl-long.238.

[5] M. d. C. Freitas и M. Mira da Silva, «GDPR Compliance in SMEs: There is much to be done,» *Journal of Information Systems Engineering & Management*, т. 3, № 4, с. 30, 2018.

[6] E. Arfelt, D. Basin и S. Debois, «Monitoring the GDPR,» en, в *Computer Security – ESORICS 2019*, K. Sako, S. Schneider и P. Y. A. Ryan, ред., Cham: Springer International Publishing, 2019, с. 681—699, ISBN: 978-3-030-29959-0. DOI: 10.1007/978-3-030-29959-0_33.

[7] Wikipedia contributors, *Text mining — Wikipedia, The Free Encyclopedia*, [Online; accessed 12-June-2024], 2023. url: https://en.wikipedia.org/wiki/Text_mining.

[8] Wikipedia contributors, *Transfer learning — Wikipedia, The Free Encyclopedia*, [Online; accessed 12-June-2024], 2023. url: https://en.wikipedia.org/wiki/Transfer_learning.

[9] J. Devlin, M.-W. Chang, K. Lee и K. Toutanova, «BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,» в *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran и T. Solorio, ред., Minneapolis, Minnesota: Association for Computational Linguistics, июнь 2019, с. 4171—4186. DOI: 10.18653/v1/N19-1423. url: https://aclanthology.org/N19-1423.

[10] M. Saqr, «Is GDPR failing? a tale of the many challenges in interpretations, applications, and enforcement,» *International Journal of Health Sciences*, т. 16, № 5, с. 1—2, 2022, ISSN: 1658-3639.

[11] D. Peloquin, M. DiMaio, B. Bierer и M. Barnes, «Disruptive and avoidable: GDPR challenges to secondary research uses of data,» en, *European Journal of Human Genetics*, т. 28, № 6, с. 697—705, июнь 2020, ISSN: 1476-5438. DOI: 10.1038/s41431-020-0596-x.

[12] A.-J. Aberkane, G. Poels и S. V. Broucke, «Exploring Automated GDPR-Compliance in Requirements Engineering: A Systematic Mapping Study,» *IEEE Access*, т. 9, с. 66 542—66 559, 2021, 5 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3076921.

[13] O. A. Cejas, M. I. Azeem, S. Abualhaija и L. C. Briand, «NLP-Based Automated Compliance Checking of Data Processing Agreements Against GDPR,» *IEEE Transactions on Software Engineering*, т. 49, № 9, с. 4282—4303, сент. 2023, 6 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 1939-3520. DOI: 10.1109/TSE.2023.3288901.

[14] Ž. Spalević и K. Vićentijević, «GDPR and challenges of personal data protection,» en, *The European Journal of Applied Economics*, т. 19, № 1, с. 55—65, 2022, ISSN: 2406-2588, 2406-3215. DOI: 10.5937/EJAE19-36596.

[15] S. Sirur, J. R. Nurse и H. Webb, «Are We There Yet? Understanding the Challenges Faced in Complying with the General Data Protection Regulation (GDPR),» в *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security*, сер. MPS '18, New York, NY, USA: Association for Computing Machinery, янв. 2018, с. 88—95, ISBN: 978-1-4503-5988-7. DOI: 10.1145/3267357.3267368. url: https://doi.org/10.1145/3267357.3267368.

[16] R. E. Hamdani, M. Mustapha, D. R. Amariles, A. Troussel, S. Meeùs и K. Krasnashchok, «A combined rule-based and machine learning approach for automated GDPR compliance checking,» en, в *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, São Paulo Brazil: ACM, июнь 2021, с. 40—49, ISBN: 978-1-4503-8526-8. DOI: 10.1145/3462757.3466081. url: https://dl.acm.org/doi/10.1145/3462757.3466081.

[17] A. Nazarenko, F. Lévy и A. Wyner, «A Pragmatic Approach to Semantic Annotation for Search of Legal Texts – An Experiment on GDPR,» E. Schweighofer, ред., 3 citations (Semantic Scholar/DOI) [2023-12-08] Book Title: Frontiers in Artificial Intelligence and Applications DOI: 10.3233/FAIA210313, IOS Press, дек. 2021, ISBN: 978-1-64368-252-5. DOI: 10.3233/FAIA210313. url: https://ebooks.iospress.nl/doi/10.3233/FAIA210313.

[18] P. A. Bonatti, S. Kirrane, I. M. Petrova и L. Sauro, «Machine Understandable Policies and GDPR Compliance Checking,» en, *KI - Künstliche Intelligenz*, т. 34, № 3, с. 303—315, сент. 2020, 27 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 1610-1987. DOI: 10.1007/s13218-020-00677-4.

[19] M. Galle, A. Christofi и H. Elsahar, «The Case for a GDPR-specific Annotated Dataset of Privacy Policies,» en,

[20] O. Amaral, S. Abualhaija, M. Sabetzadeh и L. Briand, «A Model-based Conceptualization of Requirements for Compliance Checking of Data Processing against GDPR,» в *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, 2 citations

(Semantic Scholar/DOI) [2023-12-08], сент. 2021, с. 16—20. DOI: 10.1109/REW53955.2021.00009. url: https://ieeexplore.ieee.org/document/9582337.

[21] Z. S. Li, C. M. Werner, N. A. Ernst и D. Damian, «GDPR Compliance in the Context of Continuous Integration,» *ArXiv*, февр. 2020. url: https://www.semanticscholar.org/paper/71e16573d39360b98306b3bfa5482c10b4e73746.

[22] K. Mori, T. Nagai, Y. Takata и M. Kamizono, «Analysis of Privacy Compliance by Classifying Multiple Policies on the Web,» en, в *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*, Los Alamitos, CA, USA: IEEE, июнь 2022, с. 1734—1741, ISBN: 978-1-66548-810-5. DOI: 10.1109/COMPSAC54236.2022.00276. url: https://ieeexplore.ieee.org/document/9842614/.

[23] A. Qamar, T. Javed и M. Beg, *Detecting Compliance of Privacy Policies with Data Protection Laws*. февр. 2021.

[24] S. Sousa и R. Kern, «How to keep text private? A systematic review of deep learning methods for privacy-preserving natural language processing,» en, *Artificial Intelligence Review*, т. 56, № 2, с. 1427—1492, февр. 2023, 12 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 1573-7462. DOI: 10.1007/s10462-022-10204-6.

[25] M. Srinath, S. Wilson и C. L. Giles, «Privacy at Scale: Introducing the PrivaSeer Corpus of Web Privacy Policies,» № arXiv:2004.11131, апр. 2020, 28 citations (Semantic Scholar/arXiv) [2023-12-08] arXiv:2004.11131 [cs]. DOI: 10.48550/arXiv.2004.11131. url: http://arxiv.org/abs/2004.11131.

[26] P. Silva, C. Gonçalves, C. Godinho, N. Antunes и M. Curado, «Using natural language processing to detect privacy violations in online contracts,» в *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, сер. SAC '20, 12 citations (Semantic Scholar/DOI) [2023-12-08], New York, NY, USA: Association for Computing Machinery, март 2020, с. 1305—1307, ISBN: 978-1-4503-6866-7. DOI: 10.1145/3341105.3375774. url: https://doi.org/10.1145/3341105.3375774.

[27] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin и K. Aberer, «Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning,» en, 2018, с. 531—548, ISBN: 978-1-939133-04-5. url: https://www.usenix.org/conference/usenixsecurity18/presentation/harkous.

[28] V. Leone и L. Di Caro, «The Role of Vocabulary Mediation to Discover and Represent Relevant Information in Privacy Policies,» en, в *Frontiers in Artificial Intelligence and Applications*, S. Villata, J. Harašta и P. Křemen, ред. IOS Press, дек. 2020, ISBN: 978-1-64368-150-4. DOI: 10.3233/FAIA200851. url: http://ebooks.iospress.nl/doi/10.3233/FAIA200851.

[29] N. M. Müller, D. Kowatsch, P. Debus, D. Mirdita и K. Böttinger, «On GDPR Compliance of Companies' Privacy Policies,» en, в *Text, Speech, and Dialogue*, K. Ekštein, ред., сер. Lecture Notes in Computer Science, 12 citations (Semantic Scholar/DOI) [2023-12-08], Cham: Springer International Publishing, 2019, с. 151—159, ISBN: 978-3-030-27947-9. DOI: 10.1007/978-3-030-27947-9_13.

[30] H. T. Alattas, F. M. Almassary, N. R. AlMahasheer и др., «Extract Compliance-Related Evidence Using Machine Learning,» в *2022*

14th International Conference on Computational Intelligence and Communication Networks (CICN), 0 citations (Semantic Scholar/DOI) [2023-12-08], дек. 2022, с. 537—542. DOI: 10.1109/CICN56167.2022.10008324. url: https://ieeexplore.ieee.org/abstract/document/10008324.

[31] D. Torre, S. Abualhaija, M. Sabetzadeh и др., «An AI-assisted Approach for Checking the Completeness of Privacy Policies Against GDPR,» в *2020 IEEE 28th International Requirements Engineering Conference (RE)*, 33 citations (Semantic Scholar/DOI) [2023-12-08], авг. 2020, с. 136—146. DOI: 10.1109/RE48521.2020.00025. url: https://ieeexplore.ieee.org/abstract/document/9218152.

[32] S. Arora, H. Hosseini, C. Utz и др., «A Tale of Two Regulatory Regimes: Creation and Analysis of a Bilingual Privacy Policy Corpus,» в *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, Marseille, France: European Language Resources Association, июнь 2022, с. 5460—5472. url: https://aclanthology.org/2022.lrec-1.585.

[33] M. Bokaie Hosseini, P. K C, I. Reyes и S. Egelman, «Identifying and Classifying Third-party Entities in Natural Language Privacy Policies,» в *Proceedings of the Second Workshop on Privacy in NLP*, O. Feyisetan, S. Ghanavati, S. Malmasi и P. Thaine, ред., 7 citations (Semantic Scholar/DOI) [2023-12-08], Online: Association for Computational Linguistics, нояб. 2020, с. 18—27. DOI: 10.18653/v1/2020.privatenlp-1.3. url: https://aclanthology.org/2020.privatenlp-1.3.

[34] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin и K. Aberer, «Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning,» en,

[35] Y. Ling, K. Wang, G. Bai, H. Wang и J. S. Dong, «Are they Toeing the Line? Diagnosing Privacy Compliance Violations among Browser Extensions,» в *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering*, сер. ASE '22, 7 citations (Semantic Scholar/DOI) [2023-12-08], New York, NY, USA: Association for Computing Machinery, янв. 2023, с. 1—12, ISBN: 978-1-4503-9475-8. DOI: 10.1145/3551349.3560436. url: https://dl.acm.org/doi/10.1145/3551349.3560436.

[36] D. Sánchez, A. Viejo и M. Batet, «Automatic Assessment of Privacy Policies under the GDPR,» en, *Applied Sciences*, т. 11, № 4, с. 1762, февр. 2021, 10 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 2076-3417. DOI: 10.3390/app11041762.

[37] P. Silva, C. Gonçalves, C. Godinho, N. Antunes и M. Curado, «Using NLP and Machine Learning to Detect Data Privacy Violations,» в *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 14 citations (Semantic Scholar/DOI) [2023-12-08], июль 2020, с. 972—977. DOI: 10.1109/INFOCOMWKSHPS50562.2020.9162683. url: https://ieeexplore.ieee.org/abstract/document/9162683.

[38] O. Amaral, S. Abualhaija, D. Torre, M. Sabetzadeh и L. C. Briand, «AI-Enabled Automation for Completeness Checking of Privacy Policies,» *IEEE Transactions on Software Engineering*, т. 48, № 11, с. 4647—4674, нояб. 2022, 12 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 1939-3520. DOI: 10.1109/TSE.2021.3124332.

[39] T. A. Rahat, M. Long и Y. Tian, «Is Your Policy Compliant? A Deep Learning-based Empirical Study of Privacy Policies' Compliance with GDPR,» в *Proceedings of the 21st Workshop on Privacy in the Electronic Society*, сер. WPES'22, 1 citations (Semantic Scholar/DOI) [2023-12-08] 1 citations (Crossref) [2023-12-04], New York, NY, USA: Association for Computing Machinery, нояб. 2022, с. 89—102, ISBN: 978-1-4503-9873-2. DOI: 10.1145/3559613.3563195. url: https://dl.acm.org/doi/10.1145/3559613.3563195.

[40] C. Bartolini, A. Giurgiu, G. Lenzini и L. Robaldo, «A Framework to Reason about the Legal Compliance of Security Standards,» English, нояб. 2016. url: https://orbilu.uni.lu/handle/10993/28786.

[41] P. A. Bonatti, S. Kirrane, I. M. Petrova и L. Sauro, «Machine Understandable Policies and GDPR Compliance Checking,» en, *KI - Künstliche Intelligenz*, т. 34, № 3, с. 303—315, сент. 2020, 27 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 1610-1987. DOI: 10.1007/s13218-020-00677-4. url: https://doi.org/10.1007/s13218-020-00677-4 (дата обр. 20.11.2023).

[42] D. Bui, J.-M. Choi и J. Shin, «Automated Extraction and Presentation of Data Practices in Privacy Policies,» *Proceedings on Privacy Enhancing Technologies*, т. 2021, с. 88—110, апр. 2021, 31 citations (Semantic Scholar/DOI) [2023-12-08]. DOI: 10.2478/popets-2021-0019.

[43] A. Sleimi, N. Sannier, M. Sabetzadeh, L. Briand и J. Dann, «Automated Extraction of Semantic Legal Metadata using Natural Language Processing,» в *2018 IEEE 26th International Requirements Engineering Conference (RE)*, 48 citations (Semantic Scholar/DOI) [2023-12-08], авг.

2018, с. 124—135. DOI: 10.1109/RE.2018.00022. url: https://ieeexplore.ieee.org/document/8491129?denied=.

[44] A. Nazarenko, F. Lévy и A. Wyner, «A Pragmatic Approach to Semantic Annotation for Search of Legal Texts – An Experiment on GDPR,» E. Schweighofer, ред., 3 citations (Semantic Scholar/DOI) [2023-12-08] Book Title: Frontiers in Artificial Intelligence and Applications, IOS Press, дек. 2021, ISBN: 978-1-64368-252-5 978-1-64368-253-2. DOI: 10.3233/FAIA210313. url: https://ebooks.iospress.nl/doi/10.3233/FAIA210313 (дата обр. 20.11.2023).

[45] J. M. Del Alamo, D. S. Guaman, B. García и A. Diez, «A systematic mapping study on automated analysis of privacy policies,» en, *Computing*, т. 104, № 9, с. 2053—2076, сент. 2022, 10 citations (Semantic Scholar/DOI) [2023-12-08], ISSN: 1436-5057. DOI: 10.1007/s00607-022-01076-3.

[46] R. E. Hamdani, M. Mustapha, D. R. Amariles, A. Troussel, S. Meeùs и K. Krasnashchok, «A combined rule-based and machine learning approach for automated GDPR compliance checking,» в *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, сер. ICAIL '21, 14 citations (Semantic Scholar/DOI) [2023-12-08], New York, NY, USA: Association for Computing Machinery, июль 2021, с. 40—49, ISBN: 978-1-4503-8526-8. DOI: 10.1145/3462757.3466081. url: https://doi.org/10.1145/3462757.3466081.

[47] E. Poplavska, T. Norton, S. Wilson и N. Sadeh, «From Prescription to Description: Mapping the GDPR to a Privacy Policy Corpus Annotation

Scheme,» в дек. 2020, ISBN: 978-1-64368-150-4. DOI: 10.3233/FAIA200874.

[48] N. Mousavi Nejad, P. Jabat, R. Nedelchev, S. Scerri и D. Graux, «Establishing a Strong Baseline for Privacy Policy Classification,» en, в *ICT Systems Security and Privacy Protection*, M. Hölbl, K. Rannenberg и T. Welzer, ред., сер. IFIP Advances in Information and Communication Technology, 21 citations (Semantic Scholar/DOI) [2023-12-08], Cham: Springer International Publishing, 2020, с. 370—383, ISBN: 978-3-030-58201-2. DOI: 10.1007/978-3-030-58201-2_25.

[49] S. Liu, B. Zhao, R. Guo, G. Meng, F. Zhang и M. Zhang, «Have You been Properly Notified? Automatic Compliance Analysis of Privacy Policy Text with GDPR Article 13,» в *Proceedings of the Web Conference 2021*, сер. WWW '21, 24 citations (Semantic Scholar/DOI) [2023-12-08], New York, NY, USA: Association for Computing Machinery, июнь 2021, с. 2154—2164, ISBN: 978-1-4503-8312-7. DOI: 10.1145/3442381.3450022. url: https://doi.org/10.1145/3442381.3450022.

[50] J. Giner-Miguelez, A. Gómez и J. Cabot, «DataDoc Analyzer: A Tool for Analyzing the Documentation of Scientific Datasets,» en, в *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 0 citations (Semantic Scholar/DOI) [2023-12-08], Birmingham United Kingdom: ACM, окт. 2023, с. 5046—5050, ISBN: 9798400701245. DOI: 10.1145/3583780.3614737. url: https://dl.acm.org/doi/10.1145/3583780.3614737.

[51] *Usable Privacy Policy Project — usableprivacy.org*, https://www.usableprivacy.org/data, [Accessed 13-06-2024].

[52] A. Ravichander, A. W. Black, S. Wilson, T. Norton и N. Sadeh, «Question Answering for Privacy Policies: Combining Computational and Legal Perspectives,» в *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, China: Association for Computational Linguistics, нояб. 2019, с. 4949—4959. DOI: 10.18653/v1/D19-1500. url: https://www.aclweb.org/anthology/D19-1500.

[53] R. Ramanath, F. Liu, N. Sadeh и N. Smith, «Unsupervised Alignment of Privacy Policies using Hidden Markov Models,» в *Proceedings of ACL*, Association for Computational Linguistics, июнь 2014.

[54] E. Poplavska, T. B. Norton, S. Wilson и N. M. Sadeh, «From Prescription to Description: Mapping the GDPR to a Privacy Policy Corpus Annotation Scheme,» в *International Conference on Legal Knowledge and Information Systems*, 2020. url: https://api.semanticscholar.org/CorpusID:229377855.

[55] I. L. Nicolaidou и C. Georgiades, «The GDPR: New Horizons,» en, в *EU Internet Law: Regulation and Enforcement*, T.-E. Synodinou, P. Jougleux, C. Markou и T. Prastitou, ред. Cham: Springer International Publishing, 2017, с. 3—18, ISBN: 978-3-319-64955-9. DOI: 10.1007/978-3-319-64955-9_1. url: https://doi.org/10.1007/978-3-319-64955-9_1.