

Санкт-Петербургский политехнический университет Петра Великого
Физико-механический институт

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 1
«Генерация аннотации»
по дисциплине «Автоматизация научных исследований»

Выполнил
студент гр. № 5040102/50201

Худина А.А.

Преподаватель: Новиков Ф.А.

Санкт-Петербург
2025 г.

ЗАДАНИЕ

В рамках лабораторной работы необходимо:

1. Найти любую статью с готовой оригинальной аннотацией и подготовить на ее основе документ для общения с ИИ, содержащий только текст статьи без аннотации.
2. Получить сгенерированные ИИ аннотации для выбранной статьи по четырем цепочкам промптов, предложенным в презентации, с помощью разных ИИ-моделей (использовать упомянутые в презентации, также можно дополнить своими).
3. Для первого и второго промптов при необходимости использовать ряд уточняющих промптов. Если итоговый вариант кажется неудовлетворительным, можно использовать дополнительные уточнения. В случае использования дополнительных уточнений необходимо указать в выводах текст промптов и общее количество использованных уточнений.
4. Провести сравнительный анализ полученных результатов.

Критерии оценивания ответа ИИ:

1. Соответствие заданному объему
 2. Отсутствие лексических и орфографических ошибок
 3. Соответствие заданной структуре
 4. Соответствие академическому стилю написания аннотации
 5. Соответствие содержанию статьи
5. В качестве выводов предоставить следующие результаты:
 - a. Вариант, который вы могли бы порекомендовать в качестве аннотации для выбранной статьи.
 - b. Модель, текст промпта, цепочка уточняющих промптов для выбранного варианта аннотации
 - c. Обоснование выбора

Использованная статья и модель ИИ

В лабораторной работе был использован текст статьи «Применение методов корпусной лингвистики в судебной лингвистике (на примере дела Дерека Бентли)».

Ссылка на статью: <https://elibrary.ru/item.asp?id=35727493> .

В качестве нейросети были использованы DeepSeek и ChatGPT (OpenAI).

Исходная аннотация статьи:

"Предлагаемая статья посвящена рассмотрению вопросов, которые находятся в фокусе внимания исследователей, занимающихся проблемами судебной лингвистики. Автор статьи предлагает анализ преимуществ и недостатков объединения методов корпусной и судебной лингвистики. Исследуются проблемы применения методов корпусной лингвистики в судебной лингвистике на примере дела Дерека Бентли. В контексте данного случая были приведены доказательства, которые использовались в суде, и на основании которых не были допущены судебные ошибки. Особая роль в этой ситуации принадлежит корпусной лингвистике и ее методам.

Ключевые слова: корпусная лингвистика, судебная лингвистика, дело Дерека Бентли, корпус, судебная лингвистическая экспертиза."

Промпт 1.1 (DeepSeek)

Промпт:

Составь аннотацию к загруженной статье.

Ответ:

Аннотация к статье «Применение методов корпусной лингвистики в судебной лингвистике (на примере дела Дерека Бентли)»

Статья посвящена исследованию возможностей применения методов корпусной лингвистики в рамках судебно-лингвистической экспертизы. Рассматриваются как общие принципы и проблемы этого подхода, так и конкретный пример из судебной практики — дело Дерека Бентли, осужденного в 1953 году и посмертно помилованного.

Ключевые аспекты:

- Подчеркивается роль лингвистического анализа текста как источника доказательств в уголовных и гражданских делах.
- Обсуждаются специфические сложности работы судебной лингвистики: малый объем анализируемых текстов, трудности автоматического выявления иронии, сарказма и неискренности, зависимость результатов от корректности постановки задачи.
- На примере дела Дерека Бентли демонстрируется практическое применение корпусных методов. Лингвист М. Коултхард путем сопоставительного анализа частотности и синтаксических моделей использования слова *then* в корпусах текстов (показания Бентли, других свидетелей и полицейских протоколов) доказал, что признательные показания были не записью монолога Бентли, а составлены полицейским, что в итоге привело к посмертной реабилитации обвиняемого.
- Делается вывод о том, что сочетание методов корпусной лингвистики с традиционным лингвистическим анализом является мощным инструментом для выявления и исправления судебных ошибок, борьбы с клеветой и защиты несправедливо осужденных. При этом подчеркивается необходимость крайней тщательности и осторожности в работе исследователя.

Уточнение 1:

Хорошо, но есть уточнение аннотация должна содержать цель исследования, методы, основные результаты и выводы. Объем — около 50–100 слов. Используй нейтральный академический стиль.

Ответ:

Цель исследования — определить потенциал методов корпусной лингвистики для решения задач судебно-лингвистической экспертизы на примере конкретного судебного кейса.

Методы — сопоставительный анализ частотности и синтаксических моделей использования ключевых лексем (на примере слова *then*) в специально созданных корпусах текстов: показаний подозреваемого, других свидетелей и полицейских протоколов.

Основные результаты — выявлены статистически значимые различия в употреблении языковых конструкций. Установлено, что лингвистический профиль спорных показаний Дерека Бентли существенно ближе к стилю полицейских отчетов, чем к естественной речи свидетелей, что свидетельствует о редактуре текста.

Вывод — применение корпусных методов является эффективным инструментом для выявления языковых аномалий в судебных документах и может служить веским основанием для пересмотра судебных решений.

Уточнение 2:

Стало лучше, еще одно уточнение: Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, чётко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций.

Финальный ответ:

Статья посвящена применению методов корпусной лингвистики в области судебной лингвистической экспертизы. Целью работы является оценка возможностей данного подхода для выявления лингвистических аномалий в материалах судопроизводства и установления авторства текстов. На примере дела Дерека Бентли демонстрируется применение сравнительного анализа частотности и синтаксических моделей использования ключевых лексем в специально созданных корпусах текстов. В результате исследования выявлены статистические закономерности, позволившие установить несоответствие между заявлением и реальным авторством спорных показаний. Полученные данные подтверждают значимость корпусных методов как инструмента обеспечения объективности судебных решений и расширения методологического аппарата судебной лингвистики.

Промпт 1.2 (ChatGPT)

Промпт:

Составь аннотацию к загруженной статье. Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, чётко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций.

Финальный ответ:

Статья посвящена применению методов корпусной лингвистики в судебной лингвистике на примере дела Дерека Бентли. Цель исследования — определить потенциал корпусных методов для выявления лингвистических признаков, способных повлиять на судебное решение и исправление судебных ошибок. Автор анализирует возможности и ограничения компьютерных инструментов при решении задач установления авторства и интерпретации речевых актов. В работе использованы сравнительно-корпусный и частотный анализ, выявляющий лингвистические закономерности в судебных показаниях. Результаты демонстрируют значимость корпусного подхода для судебной экспертизы текста и подчеркивают его вклад в развитие методологии судебной лингвистики.

Промпт 2.1 (DeepSeek)

Промпт:

Прочитай загруженную научную статью и напиши краткую аннотацию. Аннотация должна содержать цель исследования, методы, основные результаты и выводы. Объем — около 50–100 слов. Используй нейтральный академический стиль на русском языке.

Ответ:

Аннотация

Цель исследования — продемонстрировать возможности применения методов корпусной лингвистики в судебной экспертизе для выявления языковых маркеров, позволяющих устанавливать авторство текста и выявлять следы редактуры. В качестве основного метода используется корпусный анализ, включая сравнение частотности лексики и синтаксических структур в специализированных корпусах (свидетельских и полицейских показаний). На примере дела Дерека Бентли показано, что аномально высокая частотность слова «*then*» и его нетипичная синтаксическая позиция в показаниях обвиняемого свидетельствуют об их составлении полицией. Вывод заключается в том, что корпусные методы являются эффективным инструментом для исправления судебных ошибок, однако требуют высокой точности и осторожности в интерпретации данных.

Уточнение 1:

Уточнение: Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, чётко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций.

Финальный ответ:

Аннотация

Статья посвящена применению методов корпусной лингвистики в рамках судебной лингвистической экспертизы. Цель работы — продемонстрировать потенциал данных методов для установления авторства текста и выявления фактов его редактирования на примере конкретного судебного дела. В исследовании используется корпусный анализ, включая сравнение частотности лексических единиц и синтаксических конструкций в специализированных корпусах текстов. В результате анализа показаний по делу Дерека Бентли выявлены статистически значимые языковые маркеры, свидетельствующие о неправомерном составлении протокола. Научная значимость работы заключается в разработке верифицируемого подхода к анализу малых текстов, используемых в судопроизводстве, что способствует объективизации экспертных заключений.

Промпт 2.2 (ChatGPT)

Промпт:

Составь аннотацию к загруженной статье. Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, чётко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций.

Финальный ответ:

Аннотация:

Статья посвящена применению методов корпусной лингвистики в судебной лингвистике на примере дела Дерека Бентли. Цель исследования — показать возможности корпусного анализа при установлении авторства текста и выявлении судебных ошибок. Автор рассматривает особенности использования языковых корпусов при анализе свидетельских показаний и демонстрирует, как статистические данные и частотный анализ позволяют отличить авторские тексты от подложных. В работе применены сравнительно-частотный и контекстуальный методы анализа. Результаты подтверждают эффективность корпусных методов для судебно-лингвистической экспертизы, подчёркивая их значимость в развитии доказательной базы и предотвращении судебных ошибок.

Промпт 3.1 (DeepSeek)

Промпт:

Проанализируй загруженную научную статью и создай аннотацию, соответствующую академическим стандартам. Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, четко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций.

Финальный ответ:

Аннотация

Статья посвящена применению методов корпусной лингвистики в сфере судебной лингвистической экспертизы. Цель работы — продемонстрировать потенциал этих методов для установления авторства текста и выявления языковых индикаторов, позволяющих выявлять судебные ошибки. В качестве ключевого примера рассматривается дело Дерека Бентли, где сравнительный анализ частотности и синтаксических паттернов использования лексемы *then* в корпусах свидетельских и полицейских показаний доказал факт существенной редактуры протокола. Методология включает построение специализированных корпусов и контрастивный анализ языковых особенностей. Научная значимость исследования заключается в разработке верифицируемого инструментария для объективного лингвистического анализа в правоприменительной практике.

Промпт 3.2 (ChatGPT)

Промпт:

Проанализируй загруженную научную статью и создай аннотацию, соответствующую академическим стандартам. Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, четко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций.

Финальный ответ:

Аннотация

Статья посвящена применению методов корпусной лингвистики в судебной лингвистике, рассматриваемому на примере дела Дерека Бентли. Цель исследования — выявить возможности корпусных методов при анализе текстовых доказательств и исправлении судебных ошибок. Автор ставит вопросы о достоверности авторства и языковых особенностях свидетельских показаний. Использованы методы корпусного анализа, сопоставительного и частотного исследования. Установлено, что корпусные данные позволяют выявлять структурные и частотные закономерности, указывающие на фальсификацию текстов. Научная новизна состоит в демонстрации эффективности корпусного подхода для судебно-лингвистической экспертизы и его значения в правоприменительной практике.

Промпт:

Проанализируй загруженную научную статью и создай аннотацию. Аннотация – это один из видов сокращенной формы представления научного текста. Ее назначение – привлечь внимание читателя, пробудить читательский интерес сообщением сути исследования с помощью минимального количества языковых средств. Требования к структуре и содержанию аннотации: Информативность и содержательность. Аннотация должна в обобщенном виде представлять содержание статьи. В тексте аннотации не стоит использовать общие фразы, а также указывать несущественные детали и общеизвестные положения. Предыстория (история вопроса) может быть приведена только в том случае, если она напрямую связана с раскрытием цели исследования. Оригинальность. Следует избегать прямых повторов каких-либо фрагментов работы. Название статьи не должно дублироваться в тексте аннотации. Четкость, логичность и связность изложения. Компактность. Объем аннотаций к статьям должен составлять 50–100 слов. В аннотации должны быть представлены следующие аспекты содержания статьи: предмет, тема, цель работы (указываются в том случае, если они не ясны из заглавия статьи); метод или методология (их целесообразно описывать в том случае, если они отличаются новизной или представляют интерес с точки зрения данной работы); результаты работы (предпочтение отдается новым результатам и данным долгосрочного значения, важным открытиям, выводам, которые опровергают существующие теории, а также данным, которые, по мнению автора, имеют практическое значение); выводы (они могут сопровождаться рекомендациями, оценками, предложениями, гипотезами, описанными в статье). Не рекомендуется: включать в аннотацию таблицы, рисунки, схемы, диаграммы и формулы; приводить ссылки на номера публикаций из пристатейного списка литературы; использовать сокращения и условные обозначения, кроме общеупотребительных. Если сокращения все же вводятся, то при первом употреблении необходимо дать их расшифровку.

Финальный ответ:

Аннотация

Статья посвящена применению методов корпусной лингвистики в судебной экспертизе для анализа текстов и установления лингвистических фактов. На примере дела Дерека Бентли демонстрируется, как сравнительный анализ частотности и синтаксических паттернов (например, употребления слова *then*) позволяет выявить несоответствия в свидетельских показаниях и доказать их редактирование третьей стороной. Результаты исследования подтверждают, что корпусные методы являются эффективным инструментом для выявления судебных ошибок и могут служить основанием для пересмотра дел. Делается вывод о практической значимости методологии для установления объективной истины в правоприменительной практике.

Промпт:

Проанализируй загруженную научную статью и создай аннотацию. Аннотация – это один из видов сокращенной формы представления научного текста. Ее назначение – привлечь внимание читателя, пробудить читательский интерес сообщением сути исследования с помощью минимального количества языковых средств. Требования к структуре и содержанию аннотации: Информативность и содержательность. Аннотация должна в обобщенном виде представлять содержание статьи. В тексте аннотации не стоит использовать общие фразы, а также указывать несущественные детали и общеизвестные положения. Предыстория (история вопроса) может быть приведена только в том случае, если она напрямую связана с раскрытием цели исследования. Оригинальность. Следует избегать прямых повторов каких-либо фрагментов работы. Название статьи не должно дублироваться в тексте аннотации. Четкость, логичность и связность изложения. Компактность. Объем аннотаций к статьям должен составлять 50–100 слов. В аннотации должны быть представлены следующие аспекты содержания статьи: предмет, тема, цель работы (указываются в том случае, если они не ясны из заглавия статьи); метод или методология (их целесообразно описывать в том случае, если они отличаются новизной или представляют интерес с точки зрения данной работы); результаты работы (предпочтение отдается новым результатам и данным долгосрочного значения, важным открытиям, выводам, которые опровергают существующие теории, а также данным, которые, по мнению автора, имеют практическое значение); выводы (они могут сопровождаться рекомендациями, оценками, предложениями, гипотезами, описанными в статье). Не рекомендуется: включать в аннотацию таблицы, рисунки, схемы, диаграммы и формулы; приводить ссылки на номера публикаций из пристатейного списка литературы; использовать сокращения и условные обозначения, кроме общеупотребительных. Если сокращения все же вводятся, то при первом употреблении необходимо дать их расшифровку.

Финальный ответ:

Аннотация

Статья посвящена применению методов корпусной лингвистики в судебной практике, в частности при анализе лингвистических доказательств и исправлении судебных ошибок. Автор рассматривает возможности и ограничения использования корпусных методов для выявления авторства и интерпретации речевых конструкций. На примере дела Дерека Бентли показано, как частотный и контекстуальный анализ языковых единиц позволяет установить подделку свидетельских показаний и доказать невиновность обвиняемого. Работа подчеркивает значение корпусных технологий для повышения объективности судебно-лингвистических экспертиз и необходимости внимательного, критического подхода при их применении.

Сравнительный анализ полученных результатов

Критерии:

1. Соответствие заданному объему
2. Отсутствие лексических и орфографических ошибок
3. Соответствие заданной структуре
4. Соответствие академическому стилю написания аннотации
5. Соответствие содержанию статьи

Анализ выполнен на основе предоставленных материалов. Ниже представлена сравнительная таблица с оценкой аннотаций по заданным критериям.

Таблица 1. Сводная таблица оценки аннотаций.

Критерий / Аннотация	1.1	1.2	2.1	2.2	3.1	3.2	4.1	4.2
1. Соответствие объему (50-100 слов)	10	10	10	10	10	10	10	10
2. Отсутствие ошибок	10	10	10	10	10	10	10	10
3. Соответствие структуре	9	7	9	8	9	8	8	6
4. Соответствие акад. стилю	10	10	10	10	10	10	10	10
5. Соответствие содержанию	10	9	10	10	10	10	10	9
Средний балл	9.8	9.2	9.8	9.6	9.8	9.6	9.6	9.0

ОБОСНОВАНИЕ ОЦЕНОК:

Критерий 1: Соответствие объему (50–100 слов)

- Оценка для всех: 10. Все аннотации укладываются в заданный диапазон. Их длина составляет примерно 70-90 слов.

Критерий 2: Отсутствие лексических и орфографических ошибок

- Оценка для всех: 10. Во всех аннотациях соблюдены нормы орфографии и лексики, стилистических ошибок не выявлено.

Критерий 3: Соответствие заданной структуре

- 1.1 (9): Структура выдержанна почти идеально. Есть тема, цель, методы, результаты, значимость. Четко сформулирована цель ("оценка возможностей") и ключевой вопрос (установление авторства).
- 1.2 (7): Структура присутствует, но менее четкая. Цель сформулирована расплывчато ("определить потенциал"). Описание методов ("сравнительно-корпусный и частотный анализ") дано, но ключевой исследовательский вопрос (о редактуре текста полицией) выражен неявно. Результаты описаны обобщенно.
- 2.1 (9): Отличное соответствие структуре. Четко обозначены тема, цель, методы, результаты и научная значимость. Цель сфокусирована на установлении авторства и редактуры.

- 2.2 (8): Хорошая структура. Есть тема, цель, методы, результаты и значимость. Однако цель ("показать возможности") сформулирована чуть менее конкретно, чем в 2.1.
- 3.1 (9): Структура полная и четкая. Цель, методы, результаты и новизна изложены ясно и последовательно. Хорошо выделен конкретный пример (лексема *then*).
- 3.2 (8): Структура в целом соблюдена. Однако цель сформулирована слишком широко ("выявить возможности"), а "вопросы о достоверности" не являются четко сформулированными исследовательскими вопросами.
- 4.1 (8): Структура присутствует, но акцент смещен на пример. Цель работы как таковая не сформулирована явно, она подразумевается через демонстрацию возможностей методов. Результаты и значимость указаны четко.
- 4.2 (6): Наименьшее соответствие структуре. Тема и контекст есть. Четко сформулированная цель и исследовательские вопросы отсутствуют (есть лишь "рассматривает возможности"). Методы и результаты описаны, значимость указана. Аннотация носит более описательный, а не структурно-четкий характер.

Критерий 4: Соответствие академическому стилю

- Оценка для всех: 10. Все аннотации выдержаны в строгом, безличном, научном стиле. Используется соответствующая терминология ("корпусный анализ", "верифицируемый подход", "синтаксические паттерны", "контрастивный анализ"). Личные мнения, цитаты и списки отсутствуют.

Критерий 5: Соответствие содержанию статьи

- 1.1 (10): Полностью соответствует статье. Верно передает суть дела Бентли, методов (сравнительный анализ частотности и синтаксических моделей) и выводов.
- 1.2 (9): В целом соответствует, но упоминание "интерпретации речевых актов" не является центральной темой исходной статьи и привносит незначительный диссонанс.
- 2.1 (10): Точно отражает содержание: цель, методы, ключевой пример и выводы.
- 2.2 (10): Полное соответствие содержанию статьи.
- 3.1 (10): Идеально отражает содержание, детально и точно описывая пример с лексемой *then*.
- 3.2 (10): Верно передает содержание и основные выводы.
- 4.1 (10): Корректно отражает содержание, фокусируясь на главном кейсе и выводах.
- 4.2 (9): В целом соответствует, но фраза "доказать невиновность обвиняемого" является небольшим упрощением. Лингвистический анализ доказал фальсификацию протокола, что стало ключевым аргументом для оправдания, но не доказательством невиновности как таковым.

ЗАКЛЮЧЕНИЕ

В ходе работы были сгенерированы варианты аннотаций для статьи «Применение методов корпусной лингвистики в судебной лингвистике (на примере дела Дерека Бентли)».

Как следует из сводной таблицы результатов сравнения, все аннотации получились выполненными на довольно высоком уровне. Аннотация 3.1 является лучшей и рекомендованной, имея высший балл по всем критериям. Она наиболее точно и полно следует заданной структуре, укладывается в лимит по объему и лаконично передает суть исследования.

Промпт для её генерации в DeepSeek:

"Проанализируй загруженную научную статью и создай аннотацию, соответствующую академическим стандартам. Требования к аннотации: Язык: русский, академический стиль, без субъективных оценок. Объем: 50–100 слов. Структура: необходимо кратко обозначить тему исследования и научный контекст, чётко сформулировать цель работы и ключевые исследовательские вопросы, описать, какие методы и подходы использованы, изложить основные полученные результаты, без избыточных деталей, подчеркнуть научную новизну и значимость работы. Не включай личные мнения, цитаты, таблицы или списки. Цель: получить аннотацию, которую можно использовать в журнале или репозитории научных публикаций."