

Грамматические трансформации во вторичном сгенерированном тексте:

методология и перспективы исследования

Автор: Дикая Ю.Д., магистрант 1 курса ИИЯ "Цифровая лингвистика", РГПУ им. А.И. Герцена

Научный руководитель:
доктор филологических наук, доцент,
профессор кафедры английского языка и
лингвострановедения
Березина Ольга Александровна



1797

Современные генеративные нейросетевые модели (ИИ) активно применяются для создания вторичных текстов (рефератов, конспектов). Это меняет традиционные представления о грамматических трансформациях, которые становятся результатом алгоритмической, а не сознательной обработки.

Актуальность:

Прикладная:

Возможность улучшения самих алгоритмов генерации.

Практическая:

Задача выявления языковых "отпечатков" ИИ для систем проверки академической добросовестности (антиплагиат).

Оценочная:

Необходимость лингвистических критериев для оценки качества ИИ-текстов.

Цель и задачи:

Цель: Выявить, систематизировать и описать типичные грамматические трансформации, производимые ИИ при создании вторичных текстов.

Задачи:

1. Анализ теоретических основ (трансформации, текстуальность, принципы работы ИИ).
2. Формирование цифрового корпуса "оригинал – ИИ-конспект".
3. Сравнительно-сопоставительный анализ пар текстов.
4. Систематизация трансформаций и оценка их влияния на качество текста.

Гипотеза: ИИ-инструменты при реферировании используют устойчивую систему грамматических трансформаций, направленную на максимизацию информационной плотности при сохранении связности и содержательной адекватности.

Объект:

Процесс автоматической компрессии текста (экстрактивное и абстрактное реферирование) с помощью генеративного ИИ.

Предмет:

Типология грамматических трансформаций на синтаксическом уровне, характерных для ИИ-реферирования.

Методология:

★Эмпирические методы:

- Сравнительно-сопоставительный анализ.
- Лингвистическое описание.

★Экспериментальные методы:

Генерация конспектов с помощью выбранных ИИ-инструментов по единому протоколу.

★Теоретические методы:

Анализ литературы по грамматическим трансформациям, теории текста, компьютерной лингвистике.

★Корпусные методы для выявления

статистически значимых паттернов.

Гипотеза будет проверена на соответствие этим ожидаемых трансформаций реальным алгоритмическим паттернам ИИ.

НАУЧНАЯ НОВИЗНА:

- Системный анализ синтаксических трансформаций ИИ в научных конспектах.
- Сравнительное описание «грамматики ИИ» (ChatPDF, DeepSeek, GigaChat) при реферировании.
- Связь классической теории трансформаций с практикой ИИ-генерации текста.

Материал:

корпус параллельных текстов "статья + результат".
Источник: научные статьи.
Инструменты генерации: ChatPDF, DeepSeek.

Этапы работы:

1. Теоретический этап:

Определение теоретической базы (классификация грамматических трансформаций, теория риторической структуры, классификации реферирования).

2. Эксперимент:

- Разработка критериев отбора статей для корпуса (год публикации, тематика, структура).
- Выбор и протестирование ИИ-инструментов на пилотных статьях.
- Начало формирования корпуса.

3. Анализ:

Детальный лингвистический анализ пар текстов по намеченной схеме.

Ожидаемые типы трансформаций:

Синтаксические: Замена сложноподчиненных предложений простыми, номинализация, изменение залога (активный/пассивный), опущение модальных и эмоционально-оценочных элементов.

Текстообразующие: Изменение логико-смысловых связей между предложениями, перестройка риторической структуры (например, опущение примеров, уточнений).

Практическая значимость (ожидаемая):

1. Метрики для оценки ИИ-текстов.
2. Рекомендации по детекции ИИ-контента.
3. Основа для улучшения алгоритмов реферирования.
4. Материал для курсов по цифровой грамотности.

Заключение:

Представленное исследование находится на ключевом этапе формирования методологии и корпуса для изучения грамматического устройства ИИ-текстов. Работа направлена на создание инструмента для критического анализа и улучшения технологий генерации текста, что крайне важно в эпоху их массового применения в науке и образовании.