

Реферат

ПРОГРАММНОЕ СРЕДСТВО ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССОВ КОНТЕКСТНОГО ПОИСКА И АНАЛИЗА ИНФОРМАЦИИ В ПОЛЬЗОВАТЕЛЬСКИХ ДОКУМЕНТАХ НА ОСНОВЕ НЕЙРОСЕТЕВЫХ ЯЗЫКОВЫХ МОДЕЛЕЙ С ПРИМЕНЕНИЕМ МЕТОДА ГЕНЕРАЦИИ С ДОПОЛНИТЕЛЬНОЙ ВЫБОРКОЙ: дипломная работа / А. В. Тарканов. – Гомель: ГГТУ им. П. О. Сухого, 2026. – Дипломная работа: 135 страниц, 28 рисунков, 19 таблиц, 11 источников, пять приложений.

Ключевые слова: нейросетевые языковые модели, метод *rag*, контекстный поиск, семантический анализ, векторные эмбединги, *react*, *indexeddb*, клиентская оптимизация.

Объектом разработки являются процессы контекстного поиска и интеллектуального анализа текстовой информации в пользовательских документах различных форматов.

Цель работы: разработка автономного программного средства для автоматизации семантического поиска и интеллектуального анализа данных на основе локальных нейросетевых моделей с применением технологии генерации с дополнительной выборкой (*RAG*).

В ходе написания дипломной работы был проведен аналитический обзор современных архитектур больших языковых моделей, исследованы методы векторизации текста и алгоритмы поиска по сходству.

В процессе работы реализовано программное средство, обеспечивающее полный цикл обработки документов (*PDF*, *DOCX*, *XLSX*, *TXT*) на стороне клиента. Разработаны модули семантической сегментации текста, генерации эмбедингов и векторного поиска на базе косинусного сходства. Внедрена система кэширования данных в *IndexedDB*, позволяющая мгновенно обращаться к ранее проиндексированным документам. Уникальной функцией приложения является режим параллельного сравнения ответов от двух различных моделей в реальном времени.

Разработанная система позволяет автоматизировать поиск фактов в больших массивах неструктурированных данных, обеспечивая при этом полную конфиденциальность информации за счет локальной обработки.

Студент-дипломник подтверждает, что дипломная работа выполнена самостоятельно, приведенный в дипломной работе материал объективно отражает состояние разрабатываемого объекта, пояснительная записка проверена в системе «Антиплагиат». Процент оригинальности составляет 81,39%. Все заимствованные из литературных и других источников теоретические и методологические положения сопровождаются ссылками на источники.

