

Учреждение образования  
«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ  
Кафедра интеллектуальных информационных технологий

Отчёт  
по курсу «Естественно-языковой интерфейс интеллектуальных  
систем»

Лабораторная работа №3  
«Разработка текстового корпуса, синтаксического разбора предложений»

Выполнили студенты группы 121701:	Липский Р. В. Жолнерчик И. А. Стронгин А. В.
Проверил:	Крапивин Ю.Б.

Минск 2023

**Цель работы:**

Освоить принципы разработки прикладных сервисных программ для решения задачи анализа текста естественного языка.

**Задание:**

1. Познакомиться с назначением, структурой и функциональностью, предоставляемой базовым ЛП для решения задачи автоматического синтаксического анализа ТЕЯ.

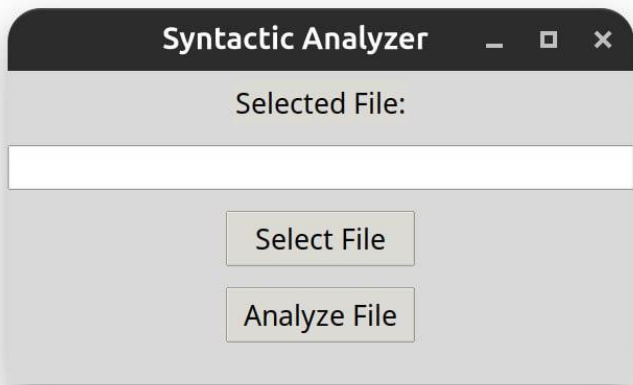
2. Закрепить навыки программирования при решении задач автоматической обработки ТЕЯ.

**Вариант: 7**

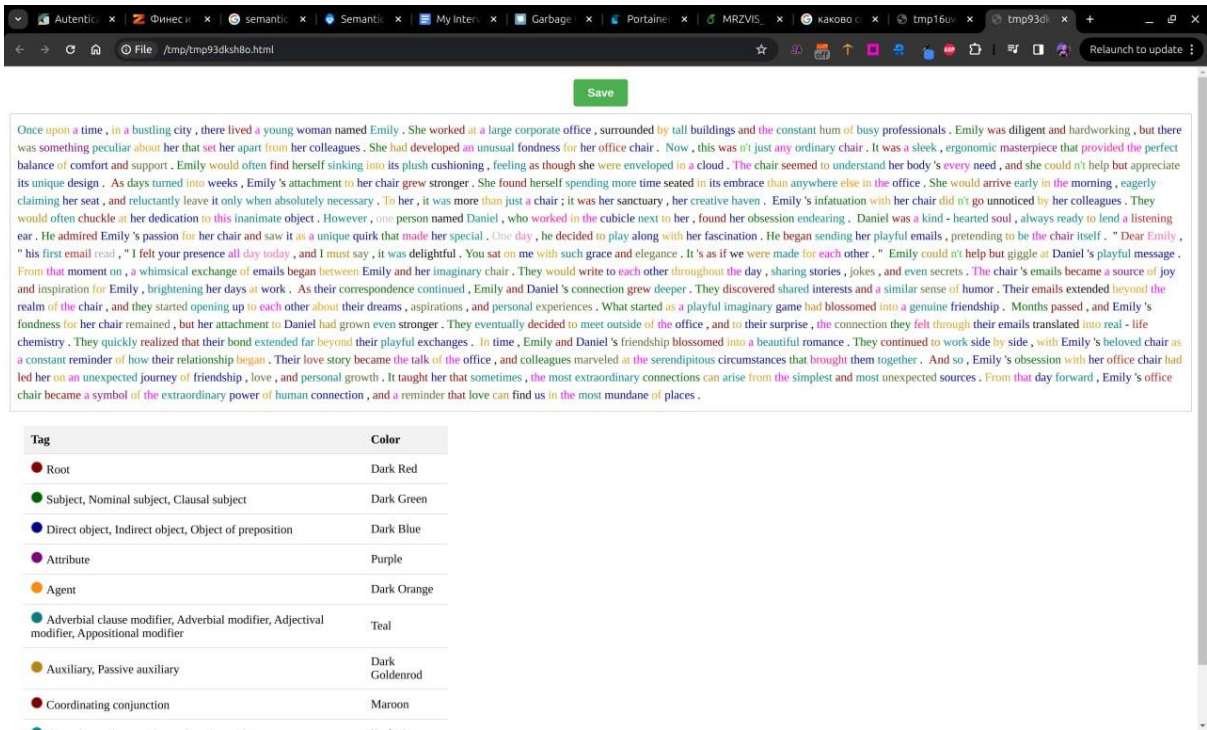
**Язык текста:** английский

**Формат входного документа:** TXT

Интерфейс:

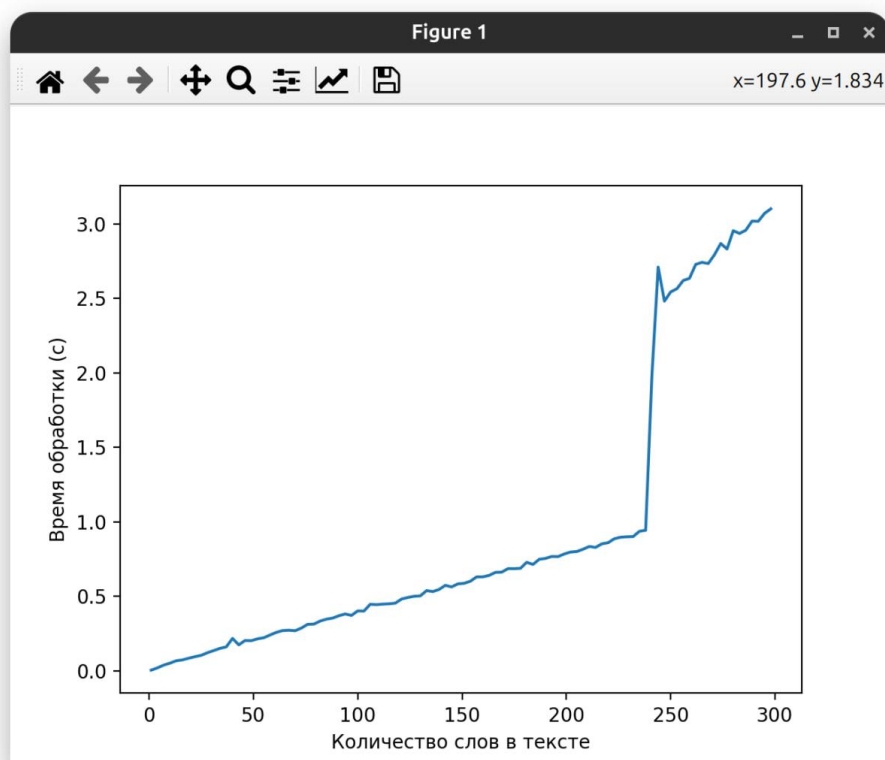


Стартовое окно выбора файла

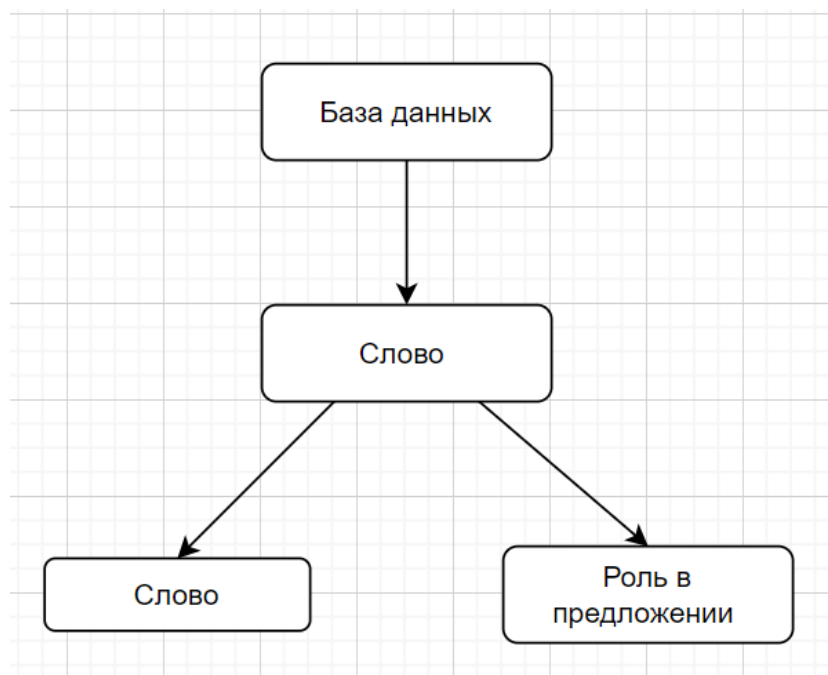


Окно вывода результата синтаксического разбора

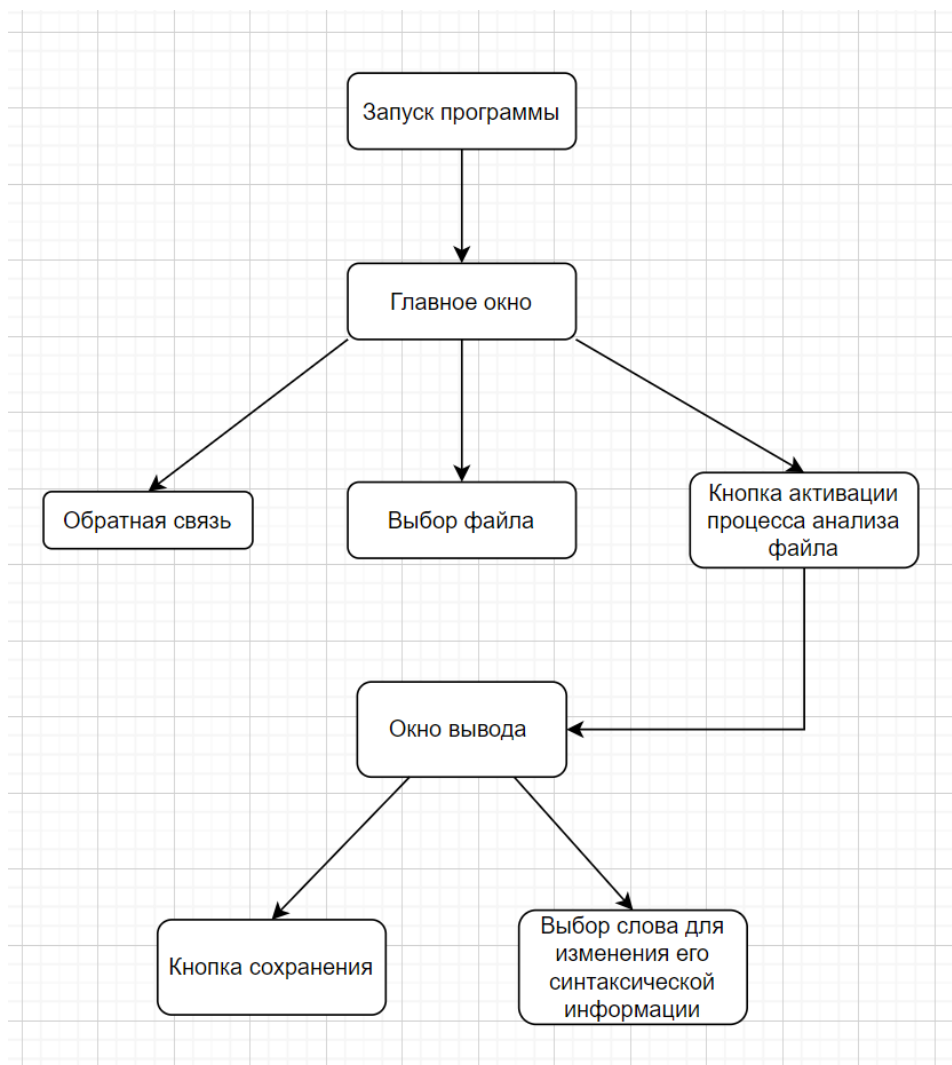
## Тест производительности:



## Структура хранения данных



### Структурная схема приложения



### Описание алгоритмов:

Разбиение текста на словоформы и получение их синтаксической информации:

Начало – Получаем файл формата TXT/RTF – Считываем текст файла – Получаем список словоформ – Получаем всю синтаксическую информацию о каждой – Генерируем размеченный HTML-текст – Отображаем HTML-файл – Конец

### Вывод:

В результате выполнения данной лабораторной работы мы изучили понятие текстового корпуса и его менеджера. Текстовый корпус представляет собой набор текстовых документов, собранных для анализа и обработки в рамках определенной задачи или исследования. Он является важным инструментом в области обработки естественного языка, машинного обучения и лингвистического анализа.