

Министерство образования Республики Беларусь

Учреждение образования
“Белорусский государственный университет
информатики и радиоэлектроники”

Факультет информационных технологий и управления
Кафедра интеллектуальных информационных технологий

Отчёт по лабораторной работе №1
по дисциплине «Естественно-языковой интерфейс интеллектуальных
систем»

Выполнил
Пашин Н.А.
Пахомов В.А.
студент группы
121701

Проверил Крапивин Ю. Б.

Минск 2024

Цель работы:

Освоить принципы разработки прикладных сервисных программ для решения задачи автоматического лексического и лексико-грамматического анализа текста естественного языка.

Задание:

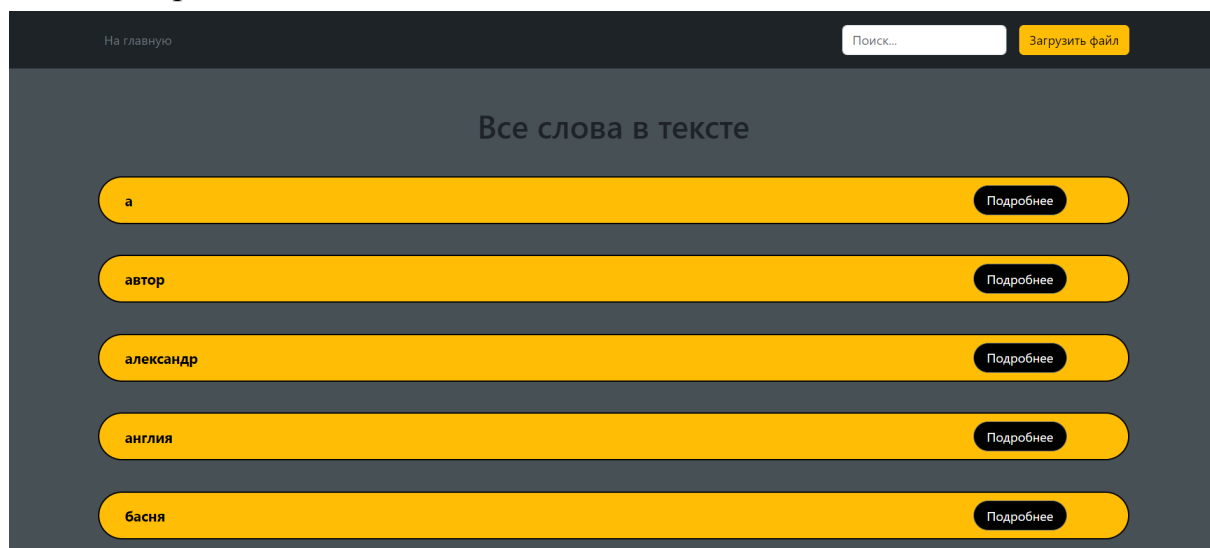
Список слов, упорядоченный по алфавиту и включающий только лексемы с дополнительно оформленными записями для образования словоформ. В этих записях должна храниться следующая информация: основа слова; часть речи; окончания слова, соотнесенные с соответствующей морфологической информацией: род, падеж, число и т.п. При работе с таким словарем должны быть обеспечены средства генерации той или иной словоформы в соответствии с введенными «правилами».

Вариант:

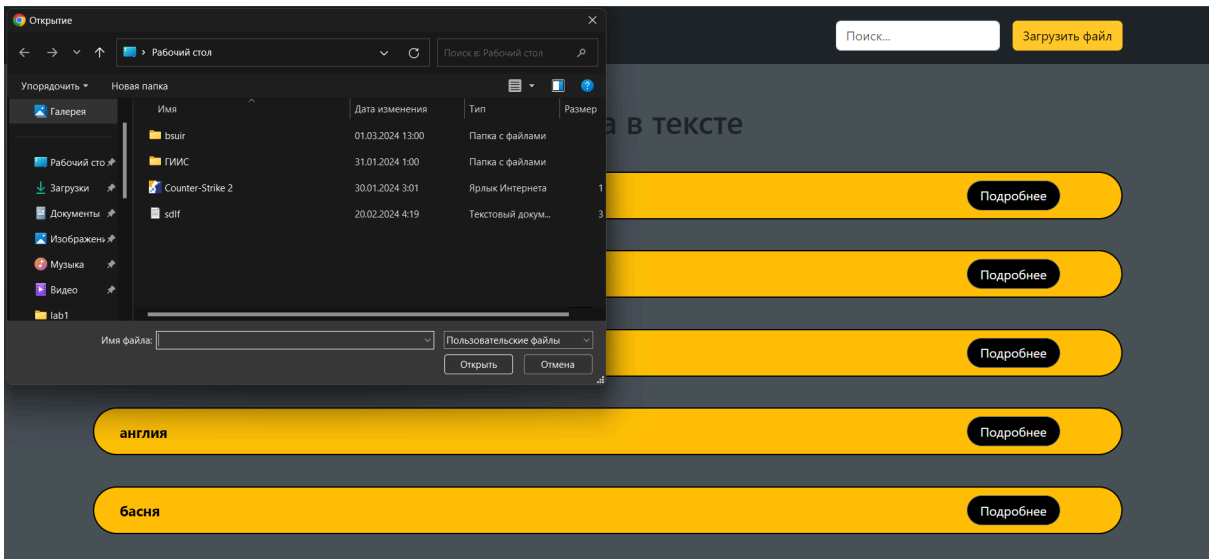
Русский	TXT, RTF	Задание 2
---------	----------	-----------

Интерфейс

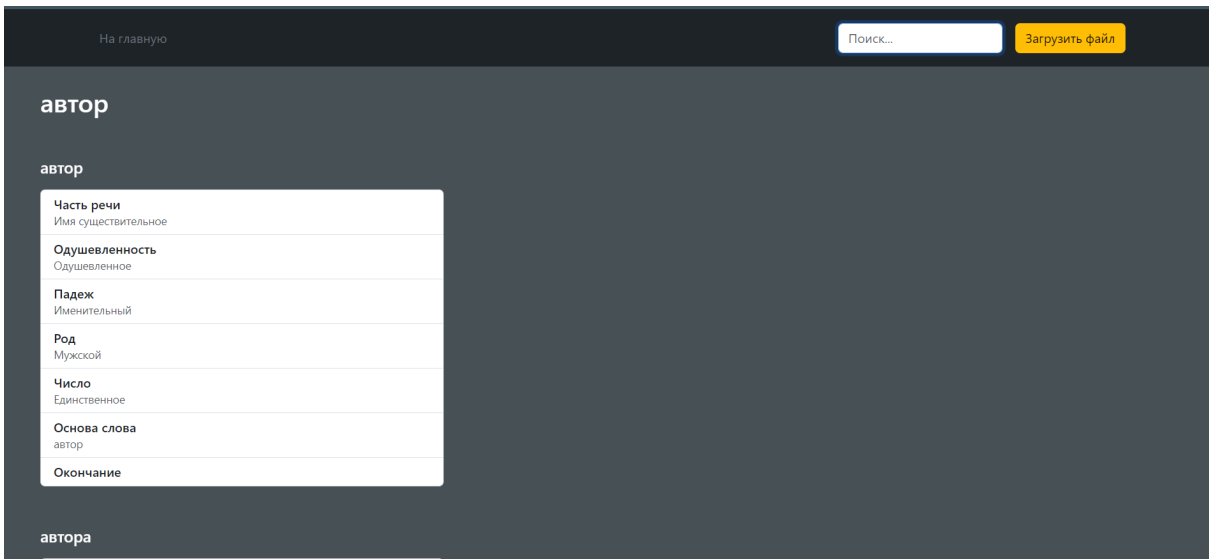
Главная страница



Окно загрузки файла



Страница одного определенного слова



Страница генерации словоформы

[На главную](#)

Поиск...

Загрузить файл

автор

Число

Падеж

Сгенерировать

Вернуться

Страница при невозможности образовать словоформу

[На главную](#)

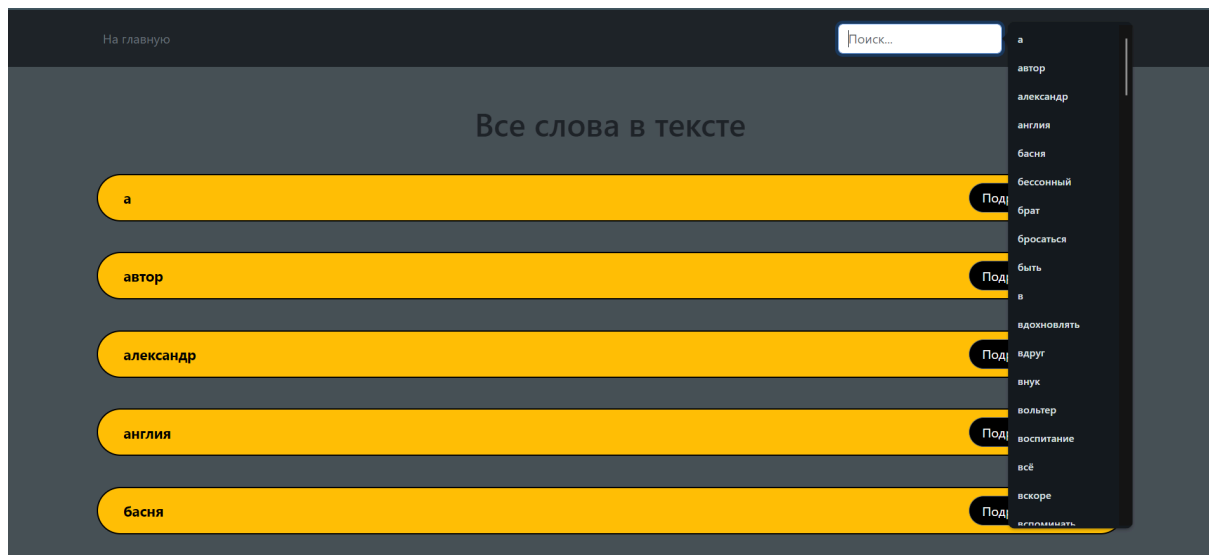
Поиск...

Загрузить файл

Для данного слова невозможно образовать словоформу

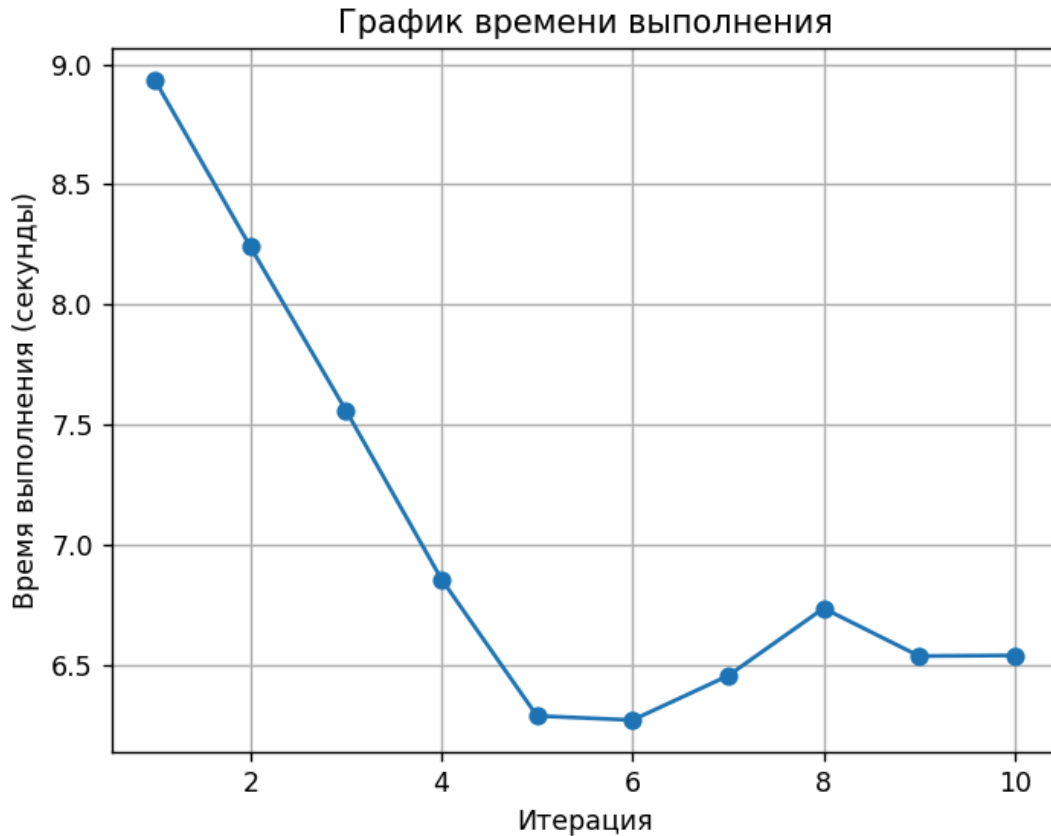
Вернуться

Окно поиска

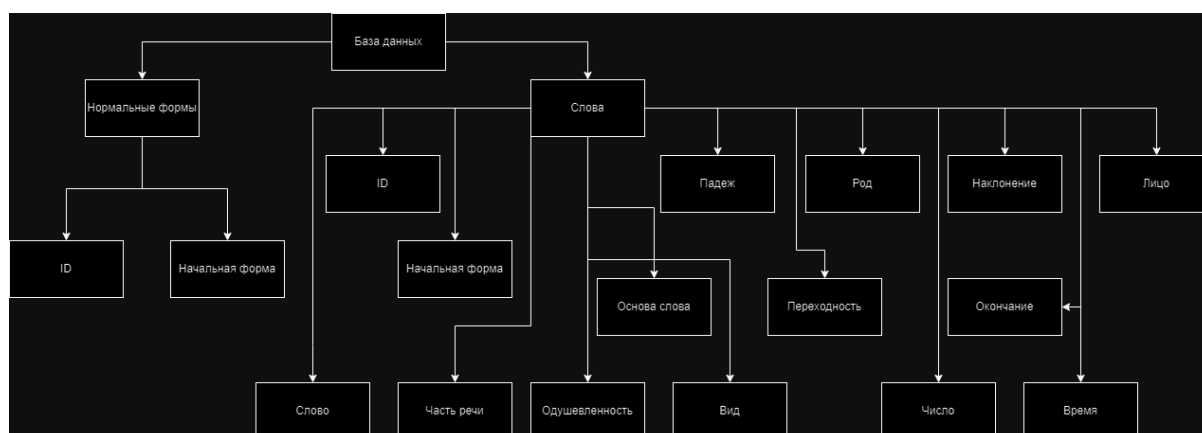


Тест производительности

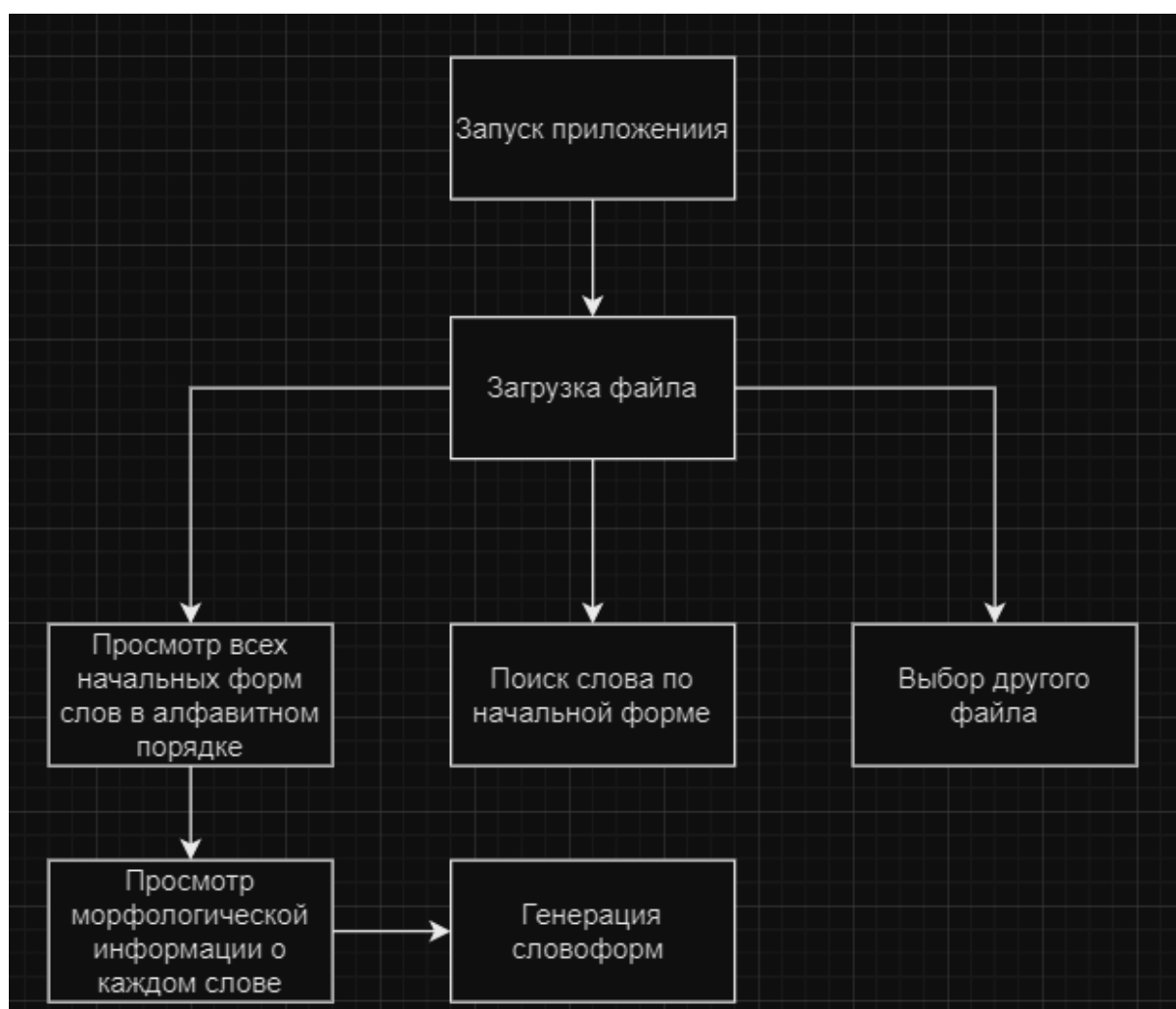
Тест производительности, приведенный на рисунке ниже, показывает время обработки текстов для разбиения на слова, нахождения морфологических признаков, основы и окончания слов и добавления данных в бд.



Структура хранения данных



Структурная схема приложения



Описание алгоритмов

Добавление слов в БД : Начало — получить файл — разобрать текст на слова — определить морфологические признаки, основу и окончание слова — добавить слово в БД — повторить последние 2 шага для оставшихся слов — Конец

Поиск по подстроке: Начало — получить начальную форму слова — найти слово в БД — получить информацию о всех словах с такой начальной формой — вывести пользователю информацию — Конец

Генерация словоформы: Начало — получить начальную форму слова — определить изменяемые морфологические признаки — вывести изменяемые морфологические признаки для данной части речи пользователю — пользователь вводит морфологические признаки — проверить могут ли быть данные морфологические признаки у слова — создать словоформу — вывести пользователю словоформу — Конец

Вывод:

Данная программа представляет собой инструмент для создания списка слов, упорядоченный по алфавиту и включающий только лексемы с дополнительно оформленными записями для образования словоформ. В этих записях хранится следующая информация: основа слова; часть речи; окончания слова, соотнесенные с соответствующей морфологической информацией: род, падеж, число и т.п. При работе с таким словарем обеспечены средства генерации той или иной словоформы в соответствии с введенными «правилами». Данная программа имеет перспективы в области анализа текстов и книг для подведения различных статистик.