Министерство образования Республики Беларусь

Учреждение образования “Белорусский государственный университет

информатики и радиоэлектроники”

Факультет информационных технологий и управления

Кафедра интеллектуальных информационных технологий

Отчёт по дисциплине

“Естественно-языковой интерфейс интеллектуальных систем”

Лабораторная работа №1

“Разработка автоматизированной системы формирования словаря

естественного языка”

Выполнил студент группы 221702: Скоробогатый А.Л

Проверил: Крапивин Ю.Б.

Минск 2025

Лабораторная работа №1

**Цель работы:**

Освоить принципы разработки прикладных сервисных программ для

решения задачи автоматического лексического и лексико-грамматического

анализа текста естественного языка.

**Задание 1, вариант 14:**

Список слов, упорядоченный по алфавиту и включающий как

лексемы, так и словоформы, с указанием частоты встречаемости каждой из форм. Для словоформ пользователю должна быть предоставлена возможность вводить дополнительную морфологическую информацию, а именно, отнесение слова к соответствующей части речи, указание рода, числа, падежа

и т.п. При этом морфологическая информация может быть оформлена как отдельная неформатированная запись, т.е. это просто текст, который пользователь может оформлять произвольным образом.

**Язык текста:**

Английский

**Формат входного документа:**

PDF

**Используемые средства разработки:**

**Tkinter** — стандартная библиотека Python для создания графических пользовательских интерфейсов (GUI)

**PyPDF2** — библиотека для извлечения текста из PDF-документов

**NLTK (Natural Language Toolkit)** — библиотека для обработки естественного языка

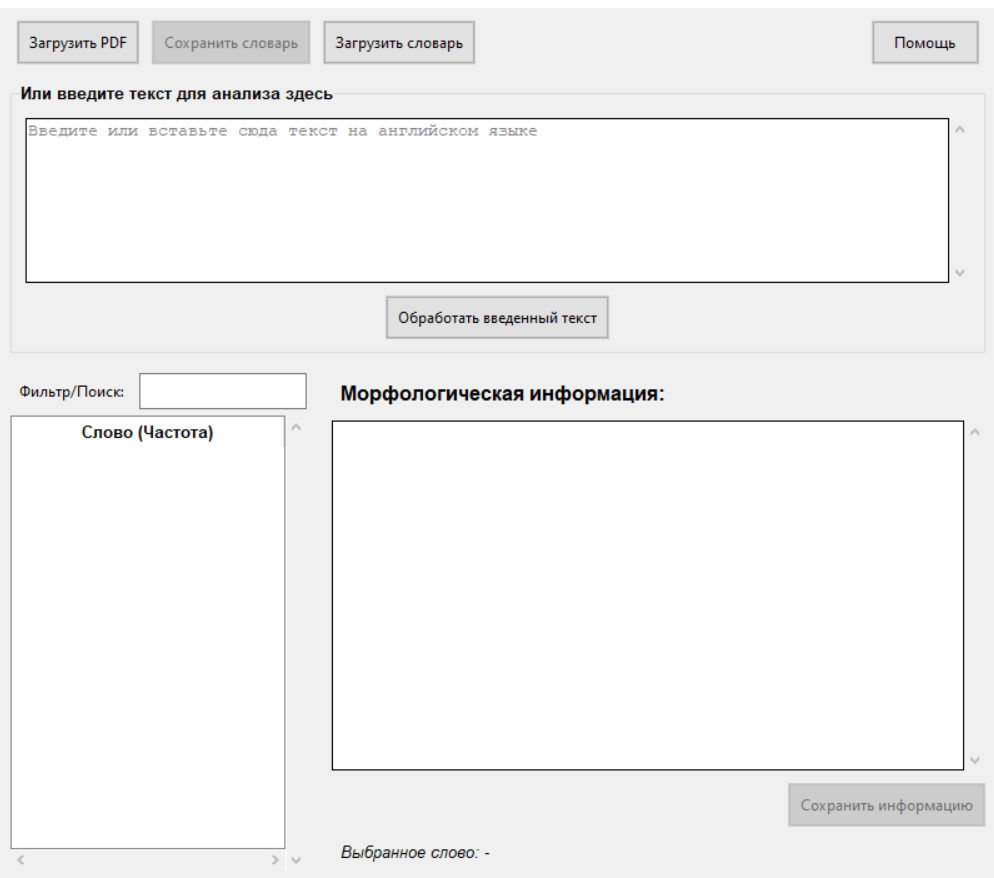
**collections.Counter** — встроенный Python-модуль для подсчёта частоты слов

**json** — стандартный модуль Python для сериализации/десериализации словаря с результатами анализа

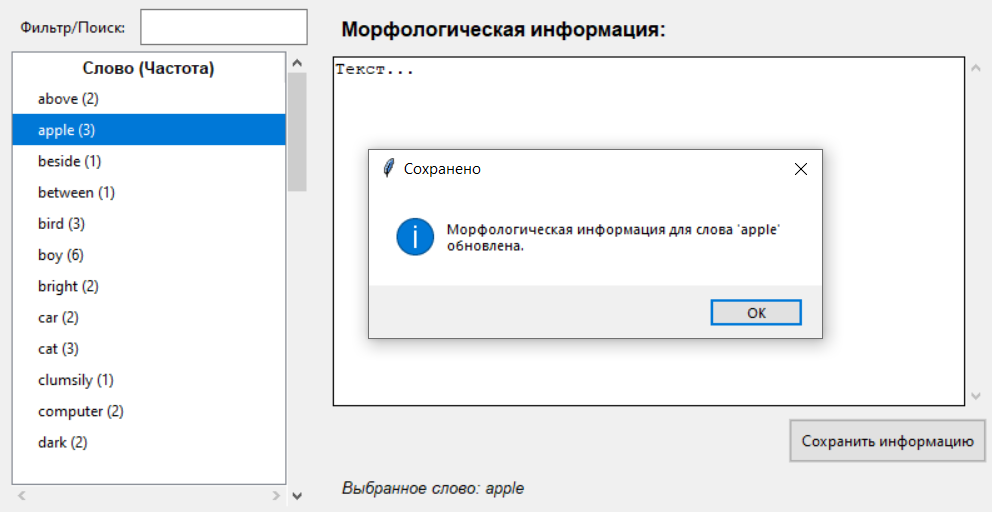
**os** — используется для работы с именами файлов и путями.

**Скриншоты разработанной системы:**

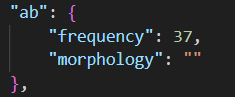
**Скриншот главного меню**



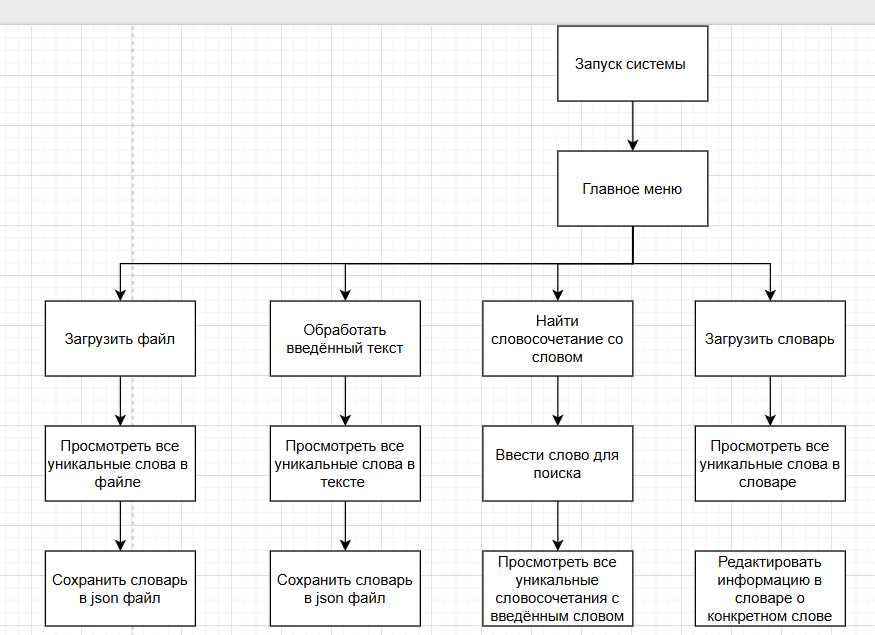
**Скриншот обновления информации**



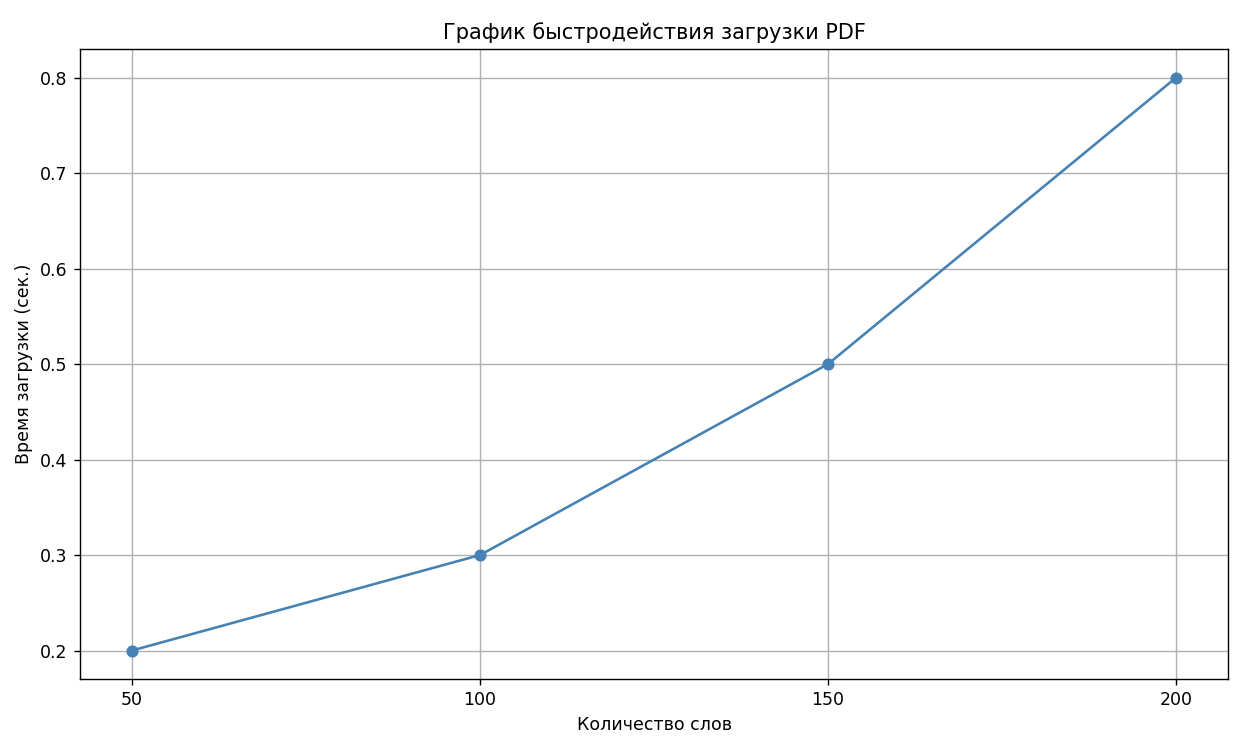
**Скриншот примера данных в словаре**



**Структурно-функциональная схема разрабатываемого приложения:**



**График оценки быстродействия приложения:**



**Выводы по работе и по перспективам развития приложения:**

Программа представляет собой автоматизированную систему формирования словаря естественного языка, представляющего собой список упорядоченных по алфавиту слов, которые могут иметь дополнительную морфологическую информацию. В качестве перспектив развития стоит добавить автоматическую морфологическую и синтаксическую разметку, оптимизировать производительность, улучшить интерфейс и добавить расширение форматов.