

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 №3, 4
по дисциплине «Естественно-языковой интерфейс интеллектуальных
систем»

Выполнил
Пашин Н.А.
студент группы
121701

Проверил Крапивин Ю. Б.

Минск 2024

Цель работы:

Освоить принципы разработки прикладных сервисных программ для решения задачи автоматического семантико-синтаксического анализа ТЕЯ.

Задание:

Познакомиться со значением, структурой и функциональностью, предоставляемой базовым для решения задачи автоматического семантико-синтаксического анализа ТЕЯ. Закрепить навыки программирования при решении задач автоматической обработки ТЕЯ.

Задача:

Предметная область: “Транспорт”.

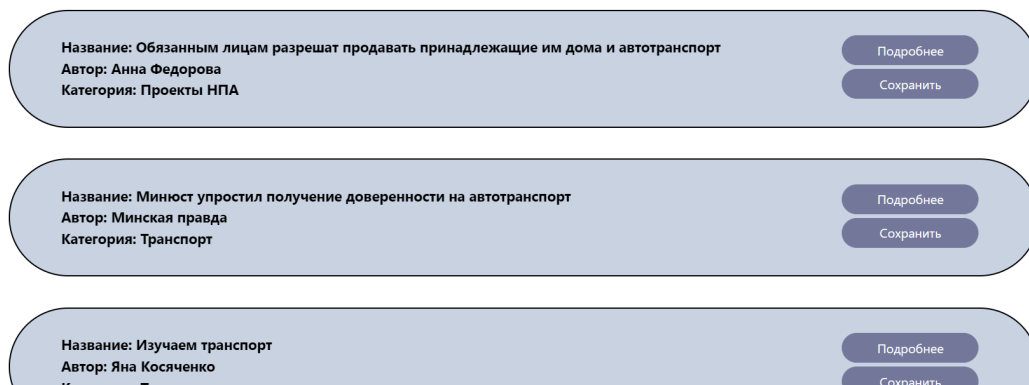
Язык: русский.

Интерфейс

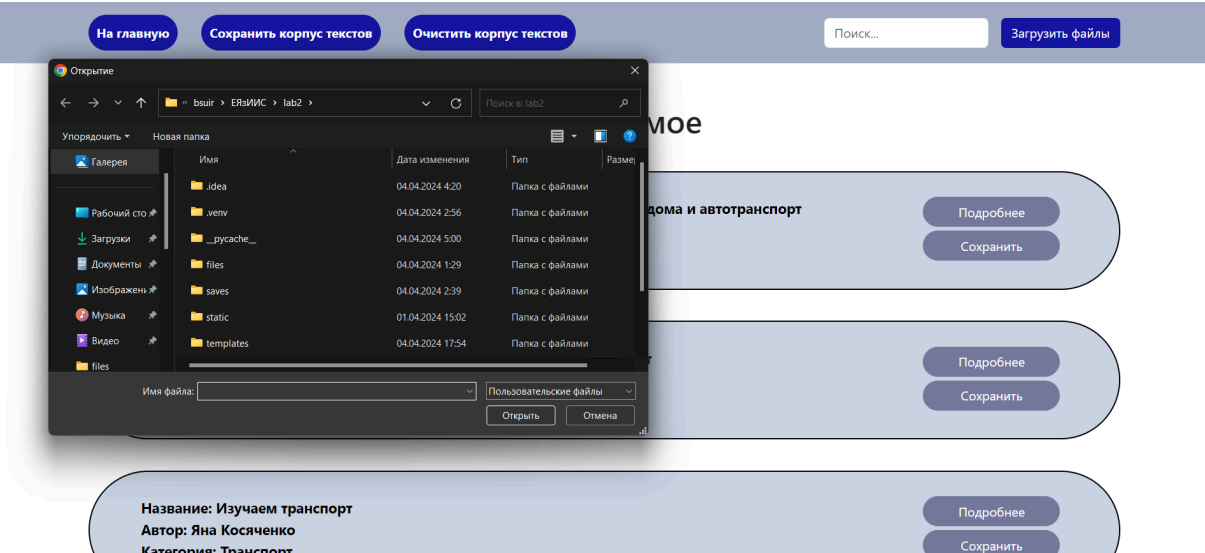
Главная страница



Содержимое



Окно загрузки файла



Страница текста



Название: Минюст упростил получение доверенности на автотранспорт

Автор: Минская правда

Категория: Транспорт

Показать текст

Показать разбор

Показать разметку

Сохранить разметку

Просмотр слов

Все виды доверенностей на машину с 13 января 2024 в Беларуси стало получить проще.

Консультант нотариальной канторы Ивьевского района Ольга Медведко прокомментировала нововведение и рассказала, как оно упростит жизнь белорусам.

По ее словам, к нотариусам часто обращаются за удостоверением доверенностей на владение, пользование и распоряжение транспортным средством, а также на временную передачу автомобиля.

Ранее нотариусу обязательно нужно было предоставить свидетельство о регистрации транспортного средства или техпаспорт.

— Не редко случалось, что при обращении к нотариусу подлинного свидетельства о регистрации транспортного средства у его владельца не оказывалось.

Название: Минюст упростил получение доверенности на автотранспорт

Автор: Минская правда

Категория: Транспорт

Показать текст

Показать разбор

Показать разметку

Сохранить разметку

Просмотр слов

В

```
<?xml version="1.0" encoding="utf-8"?><text>
<p>
<s>
<w>Все <ana lemma="всё" pos="ЧАСТ" gram="" /></w>
<w>виды <ana lemma="вид" pos="СУЩ" gram="неод,мр,мн,им" /></w>
<w>доверенностей <ana lemma="доверенность" pos="СУЩ" gram="неод,жр,мн,рд" /></w>
<w>на <ana lemma="на" pos="ПР" gram="" /></w>
<w>машину <ana lemma="машина" pos="СУЩ" gram="неод,жр,ед,вн" /></w>
<w>с <ana lemma="с" pos="ПР" gram="" /></w>
<w>13 <ana lemma="13" pos="ЧИСЛО" gram="цел" /></w>
<w>января <ana lemma="январь" pos="СУЩ" gram="неод,мр,ед,рд" /></w>
<w>2024 <ana lemma="2024" pos="ЧИСЛО" gram="цел" /></w>

```

Окно поиска

Название: Минюст упростил получение доверенности на автотранспорт

Автор: Минская правда

Категория: Транспорт

Показать текст

Показать разбор

Показать разметку

Сохранить разметку

Просмотр слов

В

```
<?xml version="1.0" encoding="utf-8"?><text>
<p>
<s>
<w>Все <ana lemma="всё" pos="ЧАСТ" gram="" /></w>
<w>виды <ana lemma="вид" pos="СУЩ" gram="неод,мр,мн,им" /></w>
<w>доверенностей <ana lemma="доверенность" pos="СУЩ" gram="неод,жр,мн,рд" /></w>
<w>на <ana lemma="на" pos="ПР" gram="" /></w>
<w>машину <ana lemma="машина" pos="СУЩ" gram="неод,жр,ед,вн" /></w>
<w>с <ana lemma="с" pos="ПР" gram="" /></w>
<w>13 <ana lemma="13" pos="ЧИСЛО" gram="цел" /></w>
<w>января <ana lemma="январь" pos="СУЩ" gram="неод,мр,ед,рд" /></w>
<w>2024 <ana lemma="2024" pos="ЧИСЛО" gram="цел" /></w>

```

Обязанным лицам разрешат продавать принадлежащие им дома и а

Изучаем транспорт

Минюст упростил получение доверенности на автотранспорт

Минюст упростил получение доверенности на автотранспорт

Окно сохранения корпуса текстов

[На главную](#)[Сохранить корпус текстов](#)[Очистить корпус текстов](#)[Загрузить файлы](#)

Введите название корпуса текстов

Сохранить

Главная страница слов

Вернуться

Все слова в тексте

- а

Подробнее
- администрация

Подробнее
- акт

Подробнее
- алименты

Подробнее

Страница одного определенного слова

Вернуться

администрация

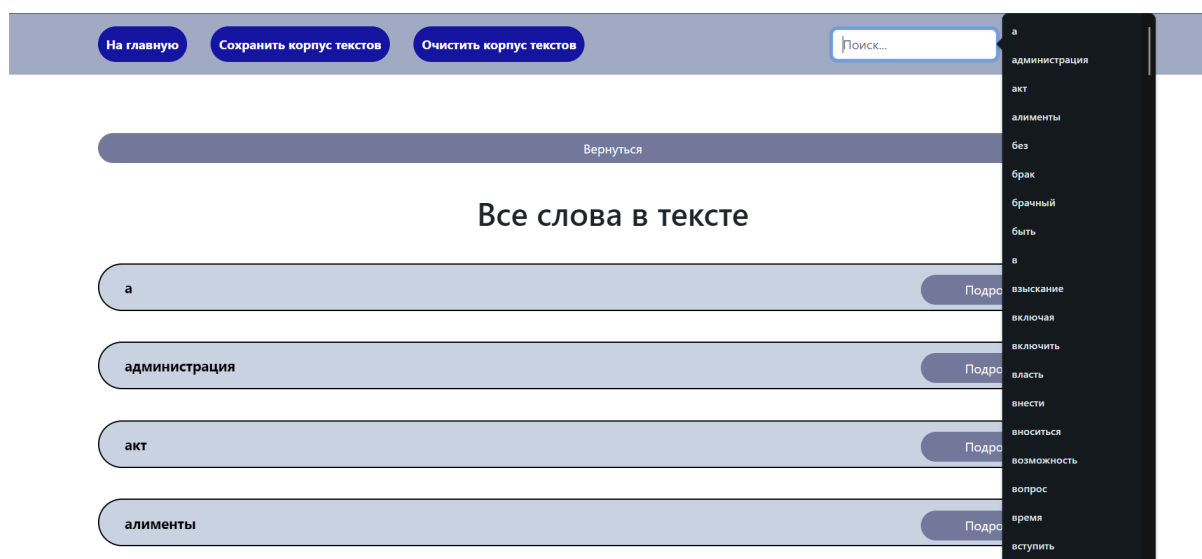
Определение

Администрация *Администрация в политике* — совокупность органов, осуществляющих функции управления в государстве или его части, а также деятельность таких органов по управлению государством или его частью. Иногда так называется определенный государственный орган. Военная администрация — орган военного управления. Государственная администрация — органы государственного управления. Местная администрация — исполнительно-распорядительный орган муниципального образования общей компетенции. Почтовая администрация — органы государства, отвечающие за регулирование, организацию и функционирование почтовой связи. Администрация президента: в США и некоторых других президентских республиках кабинет министров при главе государства, в Российской Федерации — исполнительный аппарат президента, обеспечивающий выполнение его полномочий. Региональная администрация — орган исполнительной власти общей компетенции, возглавляющий систему исполнительной власти субъекта Российской Федерации, то же, что правительство субъекта федерации. Администрация в экономике — руководство, аппарат управления фирмы, предприятия, организации, круг лиц, уполномоченных осуществлять оперативное управление, выступать в качестве юридического лица, официально представлять организацию. Временная администрация кредитной организации — специальный орган управления, назначаемый Банком России Почтовая администрация — государственный (территориальный) орган, отвечающий за регулирование, организацию и функционирование почтовой связи.

администрации

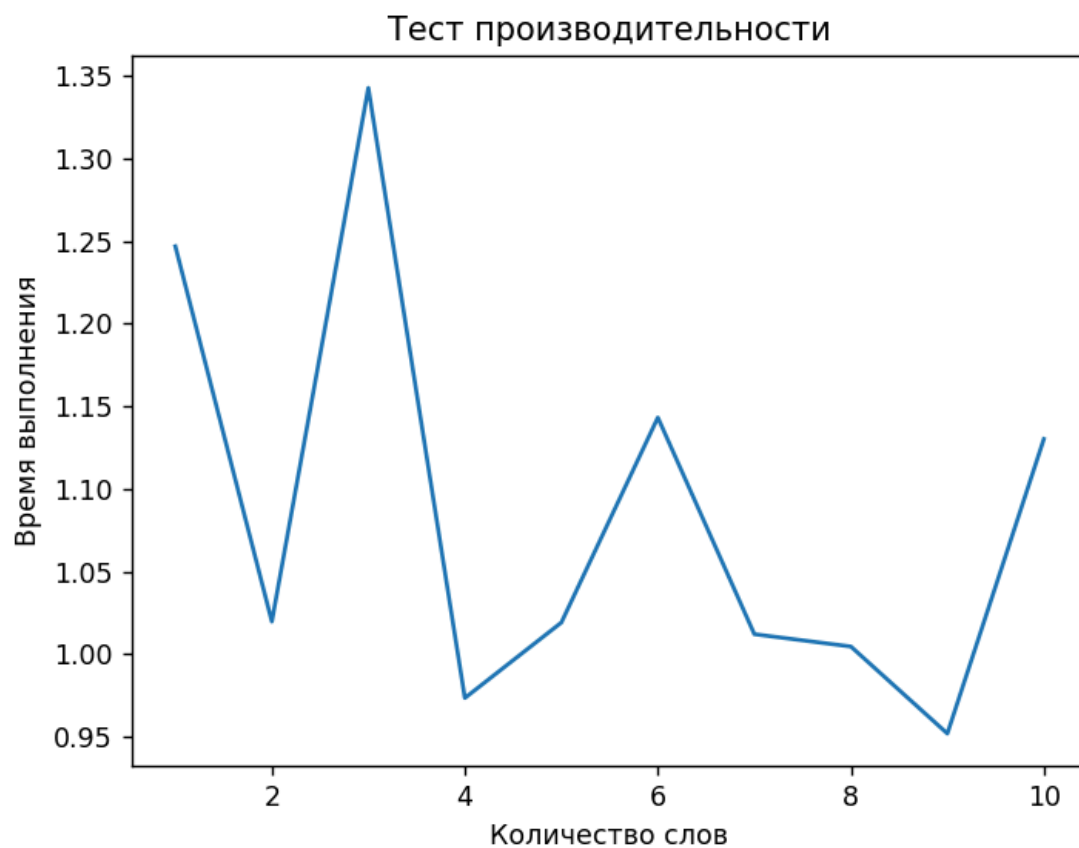
Часть речи
Имя существительное
Одушевленность

Окно поиска слов

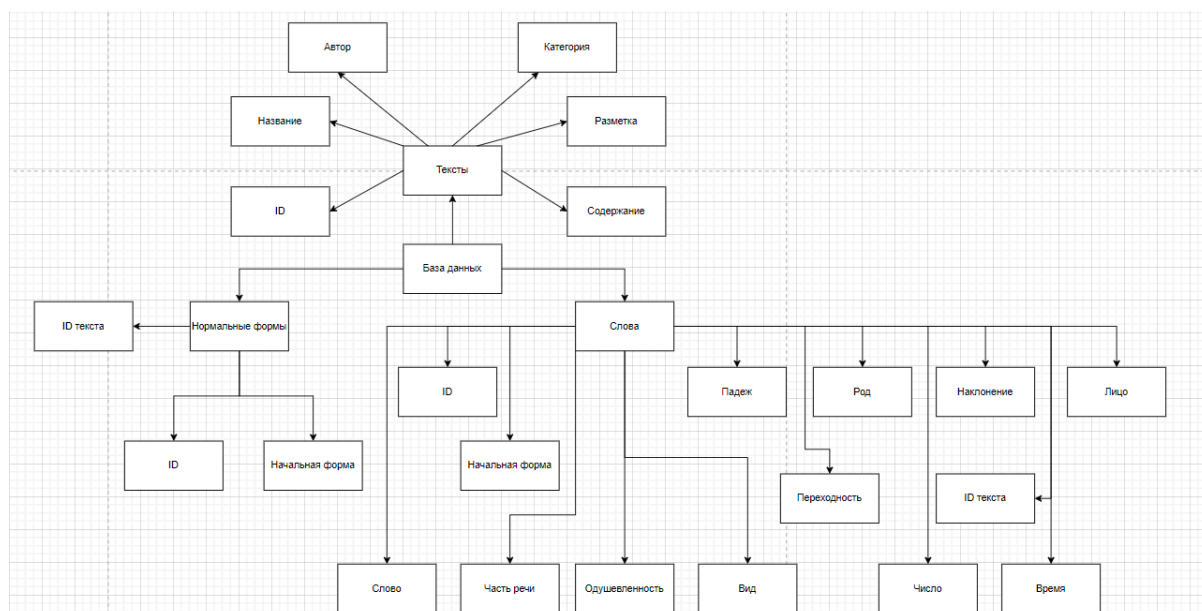


Тест производительности

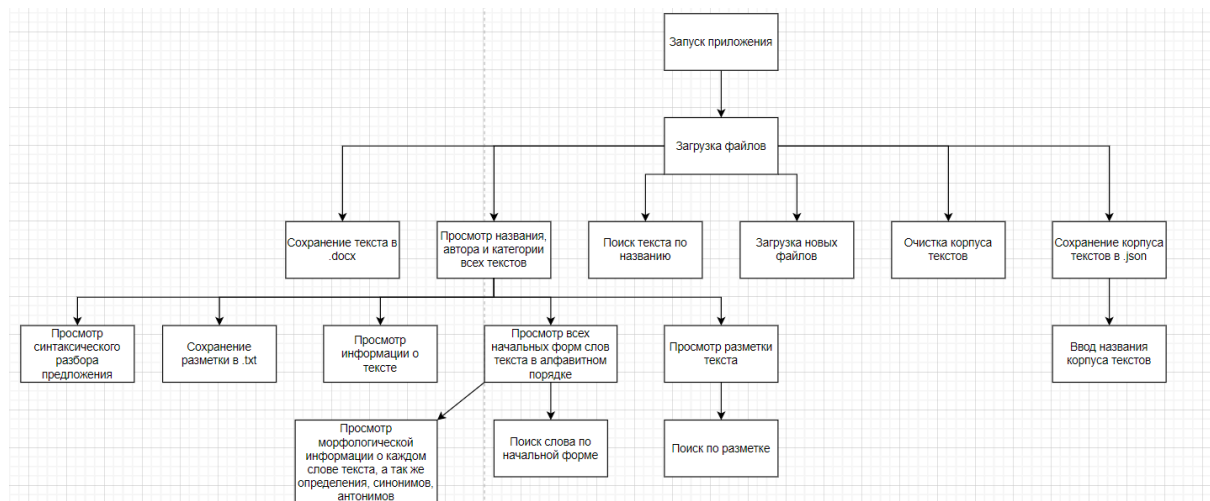
Тест производительности, приведенный на рисунке ниже, показывает время обработки для создания svg синтаксического разбора предложения в зависимости от количества слов. Конечное предложение: “Первое предложение содержит несколько слов и несет какой-то важный смысл”



Структура хранения данных



Структурная схема приложения



Описание алгоритмов

Добавление информации в БД : Начало — получить файлы — для каждого файла разобрать содержимое —определить название, автора, категорию и содержимое — создать разметку — добавить в таблицу texts — разобрать содержимое на слова — определить морфологические признаки — добавить слово в таблицу word_table — повторить последние 2 шага для оставшихся слов — Конец

Сохранение корпуса текстов: Начало — получить название корпуса текстов — получить все тексты из БД — создать файл .json — разобрать информацию и добавить в .json — выгрузить файл .json — Конец

Сохранение текста: Начало — получить текст с таким id из БД — создать файл .docx — собрать текст и добавить в файл .docx — выгрузить файл .docx — Конец

Сохранение текста: Начало — получить разметку текста с таким id текста из БД — создать файл .txt — добавить разметку в файл .txt — выгрузить файл .txt — Конец

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

Поиск по подстроке текстов: Начало — получить название текста— найти текст в БД — получить информацию о тексте с таким названием — вывести пользователю информацию — Конец

Поиск по разметке: Начало — получить искомое значение— найти в тексте — выделить при помощи `` — вывести пользователю информацию — Конец

Начало – отправляем запрос на API Википедии – получаем ответ –
Получаем тело запроса – Выделяем из него определение слова – Конец

Начало – получаем на вход слово – проверяем слово на существование его в словаре(`ru_synonyms`) синонимов и антонимов – берем все синонимы и антонимы из словаря для данного слова, если они есть - Конец

Вывод:

Лабораторная работа нам возможность ознакомиться с основами автоматического семантико-синтаксического анализа ТЕЯ, изучить структуру и функциональность базового инструментария, а также применить навыки программирования для разработки приложений в области автоматической обработки ТЕЯ. Это позволило лучше понять процессы анализа текстов и их применение в реальных ситуациях.