Государственное бюджетное общеобразовательное учреждение города Москвы «Школа №1564»

Сервис для определения основных

тематик текста

Леванова Ксения Сергеевна

10 класс ГБОУ г. Москвы №1564

Руководитель: Русаков Алексей Михайлович

Преподаватель детского технопарка «Альтаир»

**Москва, 2022**

# Оглавление

[Оглавление 1](#_Toc95851784)

[Введение 1](#_Toc95851785)

[Цель 1](#_Toc95851786)

[Задачи 1](#_Toc95851787)

[Методы и этапы 1](#_Toc95851788)

[Актуальность 1](#_Toc95851789)

[Использованные программного решения 1](#_Toc95851790)

[Обзор проектного решения 1](#_Toc95851791)

[Итоги и перспективы 1](#_Toc95851792)

[Список литературы 1](#_Toc95851793)

## Введение

В наше время все больше люди получают электронные файлы, документы и другие источники большого скопления информации, в которых зачастую сложно ориентироваться, а так как через каждого человека проходит ежедневно большое количество информации очень сложно отследить, о чем именно и когда велся диалог. Для этого используются различные технологии поиска и сортировки таких данных. В данной работе предлагается система, позволяющая проанализировать заданную информацию, для легкого ориентирования с помощью сортировки данных.

## Цель

Разработать дополнительный сервис для определения темы, чтобы облегчить и усовершенствовать его использование с помощью определения основных тематик или содержания диалога​

Также сделать интереснее и познавательнее времяпрепровождение пользователя в сфере, как различных диалогов, так и получение конкретной информации, о чем был выделенный текст.

## Задачи

1. Определить основной функционал программного сервиса​
2. Провести обзор и анализ современных технологий и выбрать наиболее пригодные​
3. Разработать прототип программного средства​
4. Определить перспективы развития проекта​
5. Создать удобный виртуальный интерфейс
6. Перенести получившийся проект в AR
7. Изучить мобильную конвертацию приложений на Android и iOS

## Методы и этапы

|  |  |
| --- | --- |
| Изучение материалов по данной теме | * Знакомство с имеющимися технологиями * Опросы специалистов * Личный опыт |
| Программирование | * Изучение фреймворка Quasar * Изучение языка Python * Изучение платформы VS Code |
| Внедрение | * Выбор самого благоприятного хостинга для реализации проекта |

## Актуальность

Проблема управления и хранения большого количества информации в огромном количестве диалогов в последнее время всё более актуальна. Для упрощения и удобного применения всей информации используются ряд технологий, как поиск по слову и разделения на разные типы вложений диалога. Технологии хорошо зарекомендованы и стандартизированы. Предлагаемая в данной работе технология развивает идею сортировки и упорядочения по основным направлениям:

1. возможность использовать данные для характеристики своего общения
2. возможность визуально увидеть на экране конкретную информацию в сжатом виде, выбранную из огромного количества информации

## Использованные программного решения

Проект создавался на языке программирования Python. Код писался в интерпретаторе Visual Studio. Создавался на платформе Quasar. После нужно выбрать нужный текст и нажать на кнопку.

## Обзор проектного решения

Выбираем текст для определения 4-ех важнейших тем и вставляем в поле. Нажимаем на кнопку «Определить тематики текста»

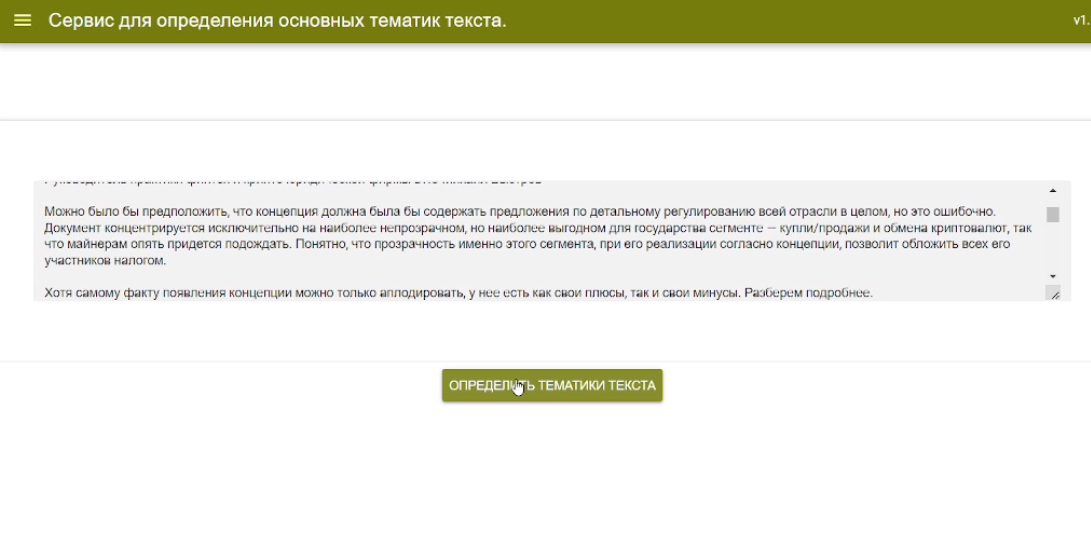


Рис. 1. Исходные данные

После обработки данных мы получаем 4 ряда слов, из которых мы можем сформулировать темы. Самые большие слова – самые значимые в данной теме.

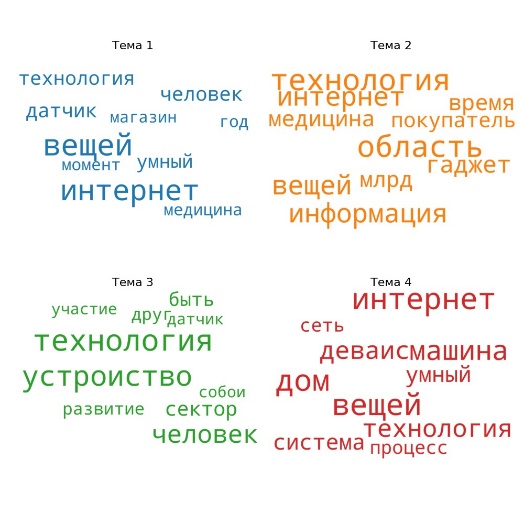


Рис. 2. Темы, полученные путем анализа текста

## Итоги и перспективы

В итоге, в проекте было реализована главная идея и задача:

- Изучены основы создания принципиальных программ;

- Описан функционал разрабатываемого сервиса;

- Разработан прототип программы и ее алгоритмы

Следующими направлениями разработки данного приложения могут быть:

1. разработка многопользовательского доступа
2. использование сервиса для других проектов
3. разработка новых функций, еще более совершенствующих и помогающих в решение данной проблемы

## Список литературы

[Электронный ресурс]. Руководство по языку программирования Python Дата обновления 05.02.2022. URL: <https://www.docs.unity3d.com/Manual/index.html>

[Учебник]. Джейми Чан. Серия: Библиотека программиста. Python. Быстрый старт. Изд-во Питер. 2021.

[Электронный ресурс] Руководство по Vue JS Прогрессивный JavaScript-фреймворк. Дата обновления 01.02.2022. URL: <https://ru.vuejs.org/index.htm>

[Электронный ресурс] Библиотека NLTK - Natural Language Toolkit Дата обновления 01.02.2022. <URL:https://www.nltk.org/>

[Электронный ресурс] Библиотека Gensim Дата обновления 01.02.2022. URL: <https://radimrehurek.com/gensim/>

[Электронный ресурс] Руководство pymorphy2 - морфологический анализатор для русского языка Дата обновления 01.02.2022. URL: <https://pymorphy2.readthedocs.io/en/0.2/user/index.html>