МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федерально бюджетное образовательное учреждение высшего образования

«Севастопольский государственный университет»

кафедра Информационных систем

Куркчи Ариф Эрнестович

Институт информационных технологий и управления в технических системах

курс 3 группа ИС/б-31-о

09.03.02 Информационные системы и технологии (уровень бакалавриата)

ОТЧЕТ

по лабораторной работе № 4

по дисциплине «Технологии обработки информации»

на тему «Исследование метода автоматического реферирования текстовой информации»

Отметка о зачете \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_

(дата)

Руководитель практикума

ст. преподаватель   Шишкевич В.Е.

(должность) (подпись) (инициалы, фамилия)

Севастополь 2017

1.Цель работы

Изучить метод автоматического реферирования текстовой информации. Научиться строить реферат на основе предложенного текстового документа.

2.Постановка задачи

Алгоритм реферирования текстов на базе статистического подхода.

1. Согласно законам Зипфа определяются характеристики исходного текста, и строится строка ключевых слов, т. е. терминов текста.

2. Фиксируется первое ключевое слово, и в реферат добавляются предложения, содержащие данное ключевое слово. При этом сохраняется порядок предложений относительно исходного текста.

3. Если достигнут требуемый размер реферата, алгоритм закончен, иначе выбирается следующее ключевое слово и повторяются шаги 2 и 3.

4. Если обработана вся строка ключевых слов, а заданный размер реферата не достигнут, из области значимых слов (рисунок 1.2) выбирается первое необработанное слово, а затем выполняется переход на шаг 2.

3.Выполнение

На рисунке 1 показан исходный текст. На рисунке 2 - ввод необходимого количества предложений для реферата. На рисунке 3 – полученный результат.

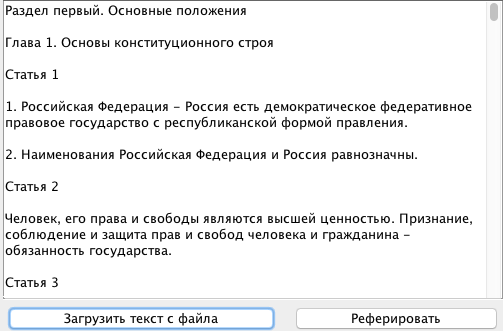


Рисунок 1 – Исходный текст

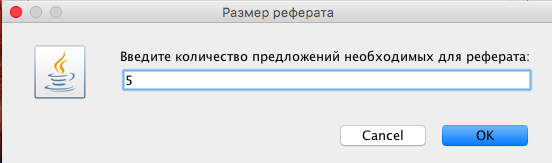


Рисунок 2 – Ввод размера реферата

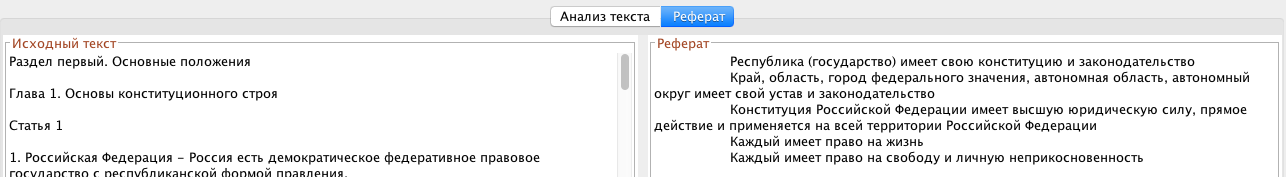


Рисунок 3 – Результат реферирования

Выводы

В ходе лабораторной работы был изучен метод автоматического реферирования текстовой информации. Получен навык строить реферат на основе предложенного текстового документа.