

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

**"БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ им. В.Г.Шухова"
(БГТУ им. В.Г.Шухова)**

Кафедра программного обеспечения вычислительной техники и автоматизированных
систем

Лабораторная работа №12

по дисциплине: "Объектно-ориентированное программирование"
по теме: "Знакомство с Python. Основные структуры данных."

Выполнил: ст. группы КБ-231
Локтионов Станислав Александрович

Проверил:
Морозов Данил Александрович

Цель работы: приобретение практических навыков создания приложений на языке Python.

Вариант 5

Задание

Разработать подсистему поиска совпадений между текстами. Первая подпрограмм будет выполнять разбор и нормализацию текста. Вторая выполнять сравнение в локальном хранилище.

Текст первого файла:

"Летний день солнечный, прогулка по парку радует душу и сердце. Вокруг цветы, слышен смех детей, все счастливы."

Текст второго файла:

"На заре осеннего дня, когда лес наполнялся мягким светом, а свежий ветер щекотал лицо, Алёна решила отправиться на прогулку. Она всегда любила это время года, когда листья меняли свои цвета на яркие оттенки золота и красного. Пройдя по узкой тропинке, девушка наслаждалась шорохом листвы под ногами и тихим пением птиц. Внезапно она остановилась, заметив маленькую полянку, залитую солнечными лучами. На ней росли грибочки, а рядом петельки дикой малины привлекали её внимание. Алёна присела, откусив сочный плод, и с улыбкой вспомнила детство, когда собирала ягоды с бабушкой. Пока она наслаждалась моментом, мысли о заботах и делах отошли на второй план. В такие моменты время кажется остановленным, и вокруг царит волшебство природы. Вдохнув глубоко свежий воздух, Алёна поняла: каждый миг жизни — это дар, который стоит ценить."

Код программы:

```
import re
import os
import string

#Подпрограмма для нормализации текста
def normalize_text(text: str) -> str:
    """Нормализация текста: нижний регистр, удаление пунктуации, лишние
    ↪ пробелов"""
    text = text.lower()
    text = re.sub(f'[{string.punctuation}]', '', text)
    text = re.sub(r'\s+', ' ', text).strip()
    return text

def read_file(filepath: str) -> str:
    """Чтение файла и нормализация текста"""
    with open(filepath, 'r', encoding='utf-8') as file:
        text = file.read()
    return normalize_text(text)

def compare_texts(text1: str, text2: str, filename1: str, filename2: str):
```

```

"""Сравнение двух текстов по общим словам"""
text1_words = set(text1.split())
text2_words = set(text2.split())

# Находим общее количество слов
common_words = text1_words.intersection(text2_words)
total_words = text1_words.union(text2_words)

if total_words: # Если есть слова в обоих текстах
    # Вычисляем процент схожести
    similarity_percentage = (len(common_words) / len(total_words)) * 100
    print(f"{filename1} и {filename2} имеют схожесть:
    ↪ {similarity_percentage:.2f}%")
else:
    print(f"{filename1} и {filename2} не имеют общих слов.")

if __name__ == "__main__":
    file1_path = '1.txt' # Путь к первому файлу
    file2_path = '2.txt' # Путь ко второму файлу

    # Чтение и нормализация текстов
    text1 = read_file(file1_path)
    text2 = read_file(file2_path)

    # Сравнение текстов
    compare_texts(text1, text2, os.path.basename(file1_path),
    ↪ os.path.basename(file2_path))

```

1.txt и 2.txt имеют схожесть: 2.52%

Рис. 1. Вывод программы

Вывод:приобрели практические навыкм создания приложений на языке Python.