

Міністерство освіти і науки України  
Національний технічний університет України «КПІ ім. Ігоря Сікорського»  
Факультет інформатики та обчислювальної техніки

Кафедра інформатики та програмної інженерії

Мультипарадигмене програмування

**ЗВІТ**

до лабораторних робіт

**Виконав**  
**студент**

Ходнєв Микита Юрійович  
(№ групи, прізвище, ім'я, по батькові )

**Прийняв**

ас. Очеретяний О. К.  
(посада, прізвище, ім'я, по батькові )

Київ 2022

## Завдання 1:

Обчислювальна задача тут тривіальна: для текстового файлу ми хочемо відобразити N (наприклад, 25) найчастіших слів і відповідну частоту їх повторення, упорядковано за зменшенням. Слід обов'язково нормалізувати використання великих літер і ігнорувати стоп-слова, як «the», «for» тощо. Щоб все було просто, ми не піклуємося про порядок слів з однаковою частотою повторень. Ця обчислювальна задача відома як **term frequency**.

Ось такий вигляд матимуть ввід і відповідно вивід результату програми:

**Input:**

```
White tigers live mostly in India  
Wild lions live mostly in Africa
```

**Output:**

```
live - 2  
mostly - 2  
africa - 1  
india - 1  
lions - 1  
tigers - 1  
white - 1  
wild - 1
```

## Опис використаних технологій

Для написання програми неможна було використовувати складні структури даних, функції, цикли. Було використано структуру. Зчитуємо з файлу в першому рядку число строк, які треба вивести. Далі по строкам з файлу читаємо текст. Оброблюємо в структуру по одному слову написаному на латиниці. Перевіряємо, чи є це слово в списку заборонених. У відповідність наводиться кількість цих слів. Далі сортуємо структуру за допомогою бульбашкового алгоритму. Виводимо слова за кількістю даних в структурі.

## Завдання 2:

Тепер, нам потрібно виконати задачу, що називається словниковим індексуванням. Для текстового файлу виведіть усі слова в алфавітному порядку разом із номерами сторінок, на яких Ці слова знаходяться. Ігноруйте всі слова, які зустрічаються більше 100 разів. Припустимо, що сторінка являє собою послідовність із 45 рядків. Наприклад, якщо взяти книгу *Pride and Prejudice*, перші кілька записів індексу будуть:

```
abatement - 89  
abhorrence - 101, 145, 152, 241, 274, 281  
abhorrent - 253  
abide - 158, 292
```

## **Опис використаних технологій**

Для написання програми аналогічно використовуємо структуру з масивом сторінок. Аналогічно до першого таску накопичуємо по строках данні та сортуємо за алфавітом.