

Министерство образования Республики Беларусь
Учреждение образования «Белорусский государственный университет
информатики и радиоэлектроники»

Факультет компьютерных систем и сетей

Кафедра информатики

Дисциплина: Операционные среды и системное программирование

ОТЧЁТ
к лабораторной работе №2
на тему

ОБРАБОТКА ТЕКСТОВОЙ ИНФОРМАЦИИ.
РЕГУЛЯРНЫЕ ВЫРАЖЕНИЯ.

Выполнил студент гр.153502 Толстой Д. В.

Проверил ассистент кафедры информатики
Гриценко Н.Ю.

Минск 2024

СОДЕРЖАНИЕ

1 Формулировка задачи	3
2 Описание функций программы.....	4
Список использованных источников	5
Приложение А (обязательное) Листинг кода.....	6

1 ФОРМУЛИРОВКА ЗАДАЧИ

Целью выполнения лабораторной работы является изучение методов и средств обработки текстовой информации, включая регулярные выражения, и использующих их утилит.

2 ОПИСАНИЕ ФУНКЦИЙ ПРОГРАММЫ

Программа обеспечивает замену строчных букв на заглавные в начале предложений, т.е. в начале документа и после точки, не находящейся внутри, например, числа, а также после знаков «!», «?».

Предложение может начинаться с новой строки (т.е. предыдущая точка может находиться в одной строке, а заменяемая строчная буква – в следующей). Пример работы скрипта представлен на рисунке 1.

```
westcrime@westcrime-Lenovo--ideapad520-15IKB:~/osisp/ЛР2$ bash script.sh text.txt
t
Это тестовый текст. Он проверяет замену строчных букв на заглавные.
В начале предложений. Также после вопросительного знака? Или восклицательного знака! 123.4 это не начало предложения. Но это начало нового предложения.
"Кавычки" тоже обрабатываются.
```

Рисунок 1 – Пример работы скрипта

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

[1] Bash manual page [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://linux.die.net/man/1/bash>

[2] Команда SED в Linux/Unix с примерами [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://habr.com/ru/companies/ruvds/articles/667490/>

ПРИЛОЖЕНИЕ А

(обязательное)

Листинг кода

```
#!/bin/bash

file_path="$1"
text=$(cat "$file_path")

# Регулярное выражение для замены строчных букв на заглавные в начале
предложений
regex='(^|[\.!?]\s+)("?[a-za-яё]) '

# Выполнение замены
result=$(echo "$text" | sed -E "s/$regex/\1\U\2/g")

echo "$result"
```