

**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»
(СПбГУТ)**

Факультет информационных систем и технологий

Кафедра безопасности информационных систем

Лабораторная работа №2-4

по дисциплине: Архитектура информационных систем

на тему: «Изучение механизма модификации создания скриптов в ОС
Windows и Linux»

Выполнил:

студент гр. ИСТ-921

Дюпин А.А.

Принял:

к.т.н., доцент кафедры ИКД

Громов В.В.

г. Санкт-Петербург

2022 г.

Цель: изучить процесс написания скриптов кодировок файлов в ОС Windows и Linux.

Ход работы

1. Работа с кодировками с помощью скрипта script.bat:

```
@echo off
chcp 866
set txt=дюпин андрей алексеевич 27.01.2001 Екатеринбург Гимназия 174
echo %txt% > 1.txt
chcp 1251
echo %txt% > 2.txt
chcp 65001
echo %txt% > 3.txt
copy *.txt result.txt
dir
type result.txt
```

2. Работа с кодировками и архивированием с помощью script.sh:

```
#!/usr/bin/bash
txt="дюпин андрей алексеевич 27.01.2001 Екатеринбург Гимназия 174"
path1="$HOME"
cd $path1
tmp_dir="lab1"
if [ -d $tmp_dir ]; then
    rm -r $tmp_dir
fi
mkdir $tmp_dir
cd $tmp_dir
echo $txt > 3.txt
iconv -f UTF-8 -t cp866 < 3.txt > 1.txt
iconv -f UTF-8 -t WINDOWS-1251 < 3.txt > 2.txt
iconv -f UTF-8 -t 8859_5 < 3.txt > 4.txt
```

```
cat *.txt > result.txt  
arj a 1.txt.arj 1.txt  
jar -cvf 2.txt.jar 2.txt  
bzip2 -k --best 3.txt  
rar a 4.txt.rar 4.txt  
zip -9 result.txt.zip result.txt  
ls -l  
cd ..  
tar -cvJSf "$tmp_dir.tar.bz2" $tmp_dir  
rm -r $tmp_dir  
echo "----->Done!<-----"
```

Вывод:

Вывод в ходе данной лабораторной работы следующий: Работа в cmd и терминале показалась довольно схожей, однако мне кажется, что в L и его дистрибутивах функции имеют более гибкий функционал из-за количества параметров. Помимо этого было выяснено, что при одних и тех же исходных данных файлы размеры получившихся .txt файлов были на 2 байта меньше при создании через терминал Ubuntu. Однако файл с расширением .bash весил больше .bat на 49 байт. При этом независимо файлы 1.txt и 2.txt совпадали по размеру, а 3.txt (UTF-8) отличался на 41 байт. Связано это с тем, что кодировка символов CP-866 и WINDOWS-1251 занимает 1 байт, а кодировка символов использует переменное количество байт от 1 до 4 (Например, символ латиницы будет занимать 1 байт, а символ кириллицы 3 байта). Однако, UTF-8 будет корректно отображаться, что у китайского пользователя, что у русского. А если вспомнить, что в 90-ые и 2000-ые года в России было как минимум три кодировки (CP-866, WINDOWS-1251, КОИ-8), то при банальном обмене сообщениями при переписке не было гарантии того, что данные будут отображаться корректно. Что уж говорить о более серьезных вещах, связанных с обменом данными. Таким образом, кодировка UTF-8 пригодна для интернационального обмена данными, но может использовать от 1 до 4 байт памяти на символ, а, например, WINDOWS-1251 кодировка локальная и подойдет для обмена данными между россиянами, но, например, не подойдет для обмена данными с Германией.