

Лабораторная работа №4.

Цель работы: получить практические навыки работы с Collections в Java, изучить работу с регулярными выражениями.

Содержание работы:

1. Изучить пакет `java.util.regex` (Классы `java.util.regex.Pattern`, `java.util.regex.Matcher`, `java.util.regex.PatternSyntaxException`)
2. Разработать программу и регулярные выражения для реализации указанного варианта задания.

Формат сдачи лабораторной работы:

1. Отчет с листингами программ и скриншотами выполнения программы.
2. Вывод по лабораторной работе.
3. Отчеты высылать на почту преподавателя с указанием темы: JAVA <Номер группы>_<ФИО>_Лаб<номер лабораторной>
4. При возникновении вопросов связываться на занятии, либо через телеграмм или почту.

Варианты заданий:

№	Задание
1	На вход программы подается произвольный файл, необходимо извлечь из него все ip адреса. В файле могут встречаться как IPv4 так и IPv6 адреса например, 127.0.0.1 или fe80:0:0:0:200:f8ff:fe21:67cf Отсортировать ip адреса по возрастанию (сначала IPv4, затем IPv6). Сохранить результат в файл.
2	На вход программы подается произвольный html-файл, необходимо извлечь из него все ссылки и посчитать их повторения. Результат сохранить в файл. Пример: На входе <html> Яндекс Яндекс карты Поиск > <p>Много много текста а подробнее можно посмотреть тут http://podrobnее.ru </p> </html> На выходе: http://ya.ru 3 http://ya.ru/maps 4 http://www.google.ru/search?hl=ru&client=firefox&hs=mpE&rls=en 1 http://podrobnее.ru 3

3	<p>На вход программы подается произвольный html-файл, необходимо извлечь из него все адреса электронной почты. Следует учесть, что email-адреса могут быть также в зоне рф т. е. Например тест@тест.рф Напротив каждого варианта почты указать число повторений. Результат сохранить в файл. Пример: На входе <html> <p> Вася пупкин vasy.pupkin@gmail.com</p> <p> Вася пупкин vasy.pupkin@yandex.com</p> </html> На выходе: vasy.pupkin@gmail.com 5 vasy.pupkin@yandex.com 1</p>
4	<p>Пользователь вводит последовательно строкой за строкой регистрационные данные <i>Фамилия Имя Отчество мобильный телефон ник дата рождения.</i> Необходимо с помощью одного регулярного выражения проверить корректность введенных данных. Вывести введенную информацию в следующем виде (для извлечения введенных данных использовать группы): Фамилия – Петров Имя - Петр Отчество – Петрович Моб. Телефон – 7915715xxxx Ник – petr.petr Дата рождения – 24.07.1989</p> <p>По завершению ввода данных пользователей, должен сохраниться файл со всеми введенными данными</p>
5	<p>Пользователь вводит регистрационные данные о сотрудниках. Необходимо проверить их на корректность. Фамилия _ может состоять из русских или английских букв, но не вперемешку и содержать знак тире. Имя _ может состоять из русских или английских букв, но не вперемешку. Мобильный телефон _ может быть записан в разных форматах:</p> <ul style="list-style-type: none"> • 791557955xx • +791557955xx • 891557955xx • 7(915)579-55-xx • 7-915-57-95-5xx • и т. д. <p>Дата рождения — может быть записана в разных форматах:</p> <ul style="list-style-type: none"> • ДД.ММ.ГГГГ • ДД.ММ.ГГ • ГГ.ГГММ.ДД • и т. д. <p>В качестве разделителя может использоваться точка, тире и двоеточие т. е. Даты перечисленные ниже корректны:</p>

	<ul style="list-style-type: none"> • 10.10.2010 • 10-10-2010 • 10:10:2010 <p>По завершению ввода данных пользователей, должен сохраниться файл со всеми введенными данными</p>
6	<p>На вход программы подается произвольный html-файл, необходимо удалить все атрибуты у всех тегов, кроме заданных пользователем.</p> <p>Вывести в консоль статистику по обработанному файлу:</p> <ul style="list-style-type: none"> - указать тег и кол-во в исходном файле из заданных пользователем - указать тег и кол-во в исходном файле у которых удалены атрибуты <p>Сохранить новый html-файл.</p>
7	<p>На вход программы подается произвольный html-файл, необходимо все ссылки <code>http://</code> не находящиеся в теге <code><a></code> обернуть тегом <code><a></code> в качестве описания использовать имя домена до первого символа <code></></code>.</p> <p>Например, есть ссылка https://www.bsu.edu.ru/bsu/resource/schedule/teachers/index.php?teacher=150016 Необходимо заменить на <code>bsu.edu.ru</code></p> <p>Вывести в консоль статистику по обработанному файлу:</p> <ul style="list-style-type: none"> - указать ссылки с тегом <code><a></code> и их количество повторений - указать ссылки без тега <code><a></code> обработанные и их повторения <p>Сохранить новый html-файл.</p>
8	<p>На вход программы подается произвольный html-файл необходимо извлечь из него значения всех атрибутов. Пример:</p> <pre>.... ТЕСТ <p align='center'>текст </p></pre> <p>В результате должно быть:</p> <pre>href http://test.ru align center</pre> <p>Сохранить результаты в файл.</p> <p>Вывести в консоль статистику по обработанному файлу:</p> <ul style="list-style-type: none"> - кол-во каждого из атрибута
9	<p>На вход программы подается произвольный текст, включающий т.н. ВВ-коды. Преобразовать все ВВ-коды в корректные html-теги.</p> <p>Пример:</p> <pre>[font="Arial"]текст[/font] [size="14px"]текст[/size] [color="#FF0000"]текст[/color] [url="http://test.ru"]ссылка[/url] ...</pre> <p>В результате должно быть:</p> <pre>текст текст текст ссылка </pre>

	<p>Учесть, что ВВ-коды могут быть вложены друг в друга, например [font="Arial"][size="14px"]часть[/size] текста [font="Times New Roman"] другим [/font] шрифтом[/font]</p> <p>Вывести в консоль статистику по обработанному файлу: - указать тип ВВ-кода и его количество повторений</p>
10	<p>На вход программе подается произвольный html-файл. Необходимо всем тегам добавить недостающие атрибуты height и width со значениями, заданными пользователем.</p> <p>Пример: </p> <p>На выходе: </p> <p>100, 200 и 300 это значения, введенные пользователем Отсортировать теги по размеру height и width по возрастанию. Сохранить новый html-файл.</p>
11	<p>На вход подается HTML файл. Найдите все ссылки вида «» в этом документе и выведите список сайтов, на документы которых он ссылается. Сайты следует выводить в алфавитном порядке, формат ссылки — см RFC 3986 (http://tools.ietf.org/html/rfc3986)</p> <p>Пример: Яндекс Яндекс карты Поиск</p> <p>На выходе: ya.ru google.ru</p>
12	<p>На вход подается текстовый файл. При помощи регулярных выражений программа разбивает текст на отдельные слова и находит частоту встречаемости отдельных слов.</p>
13	<p>На вход программы подается произвольный html-файл, необходимо удалить все атрибуты у всех тегов, кроме заданных пользователем.</p> <p>Вывести в консоль статистику по обработанному файлу: - указать тег и кол-во в исходном файле из заданных пользователем - указать тег и кол-во в исходном файле у которых удалены атрибуты Сохранить новый html-файл.</p>
14	<p>Пользователь вводит последовательно строкой за строкой регистрационные данные <i>Автомобильный номер Производитель Модель Год производства.</i> Необходимо с помощью одного регулярного выражения проверить корректность введенных данных. Вывести введенную информацию в следующем виде (для извлечения введенных данных использовать группы): Автомобильный номер – C134YP31 Производитель - Audi Модель – Q5 Год производства – 2018</p>

	По завершению ввода данных пользователей, должен сохраниться файл со всеми введенными данными
15	<p>На вход программы подается произвольный текстовый файл, необходимо извлечь из него все пароли, которые удовлетворяют условию от 8 до 15 символов с минимум одной цифрой, одной заглавной и одной строчной буквой. т. е. Например A32DjodpL.</p> <p>Напротив каждого подходящего пароля указать число повторений. Результат сохранить в файл.</p> <p>Пример: На входе A32DjodpL Авфыа Ва P232kfPfkdkfl A32DjodpL</p> <p>На выходе: A32DjodpL 2 P232kfPfkdkfl 1</p>
16	<p>На вход программы подается произвольный файл, необходимо извлечь из него все телефонные номера. Например, +79302303203 или +7 (4722) 583034</p> <p>Отсортировать телефонные номера по возрастанию. Сохранить результат в файл.</p>
17	<p>На вход программы подается произвольный html-файл, необходимо извлечь из него все ссылки и посчитать их повторения. Результат сохранить в файл.</p> <p>Пример: На входе <html> </html></p> <p>На выходе: img1.jpg 2 img2.jpg 1</p>
18	<p>На вход программы подается произвольный файл, необходимо извлечь из него все email адреса. Например, abc@gmail.com</p> <p>Отсортировать email по алфавиту. Сохранить результат в файл.</p>