

PONTIFICIA UNIVERSIDAD CATOLICA MADRE Y MAESTRA



Nombre:

Miguelson Antenor

2010-2240

Materia:

ISC-415

Programacion web

Profesor:

Ing. Carlos Camacho

Reporte sobre:

Practica 1

PROGRAMACIÓN WEB
Creación Cliente HTTP – Apache Components
Práctica #1

Objetivos:

- Trabajar con el protocolo HTTP mediante la librería Apache Components y SOAP.
- Trabajar con gestión de proyectos basado en Gradle.

Desarrollo de la práctica.

El estudiante realizará cada uno de los ejercicios planteado en dicho documento. El proyecto proceso de la realización de la práctica debe estar evidenciado en el control de versiones bajo la tecnología Git. La práctica será individual.

Forma de Entrega:

La entrega de la práctica se realizará un reporte conjuntamente con el envío la asignación al servidor de versiones, debe estar lista para ejecutar con la tarea run del plugin application.

Desarrollo

**En un proyecto basado en Gradle y utilizando la librería de Apache Components, específicamente el modulo HTTP Client 4.4 (<http://hc.apache.org/>) o la librería JSOAP (<http://jsoup.org/>), estaremos interpretador la trama HTML y aplicando las siguientes:
Crea un programa por consola que pida por la entrada estándar una URL valida, una vez consultada realice las siguientes operaciones:**

- a) Indicar la cantidad de lineas del recurso retornado.**
- b) Indicar la cantidad de párrafos (p) que contiene el documento HTML.**
- c) Indicar la cantidad de imágenes (img) dentro de los párrafos que contiene el archivo HTML.**
- d) indicar la cantidad de formularios (form) que contiene el HTML por categorizando por el método implementado POST o GET.**
- e) Para cada formulario mostrar los campos del tipo input y su respectivo tipo que contiene en el documento HTML.**
- f) Para cada formulario “parseado”, identificar que el método de envío del formulario sea POST y enviar una petición al servidor con el parámetro llamado asignatura y valor practica1 y un header llamado matricula con el valor correspondiente a matrícula asignada. Debe mostrar la respuesta por la salida estándar.**

Tecnología utilizada

Gradle es una herramienta de automatización de compilación para el desarrollo de software en varios idiomas. Controla el proceso de desarrollo en las tareas de compilación y empaquetado para pruebas, implementación y publicación. Los lenguajes admitidos incluyen Java, C/C++ y JavaScript.

jsoup es una biblioteca Java de código abierto diseñada para analizar, extraer y manipular datos almacenados en documentos HTML.

IntelliJ IDEA es un entorno de desarrollo integrado escrito en Java para desarrollar software informático escrito en Java, Kotlin, Groovy y otros lenguajes basados en JAR. Está desarrollado por JetBrains y está disponible como una edición comunitaria con licencia de Apache 2 y en una edición comercial propietaria.

GitHub, Inc. es un proveedor de alojamiento de Internet para el desarrollo de software y el control de versiones mediante Git. Ofrece el control de versiones distribuidas y la funcionalidad de gestión de código fuente de Git, además de sus propias características.

Para realizar este trabajo yo tenía que tomar en cuenta la estructura del documento HTML

```
<!DOCTYPE html>
<html>
  <head>

    <title>Page Title</title>

  </head>

  <body>

    <h1>This is a Heading</h1>
    <p>This is a paragraph.</p>

  </body>

</html>
```

El documento principal está compuesto y estructurado de balística llamado tag.

<html> es la entrada del documento y el <head> es la cabeza donde se puede insertar meta informaciones tal como la idioma, tambien link de CDN para relevante herramientas par el buen funcionamiento de la página y el nombre de la página <title>.

<body> donde se insertará todo el contenido de la página.

`<h1>` información que tiene que ver con el título de un sujeto son 6 en total.

`<p>` para insertar párrafos.

Dentro de estos tag van algunos datos llamado atributo con su valor: id, class, method en caso de un form, etc.

Con la librería Jsoup creó un documento para luego manipularlo como el DOM.

Huize principalmente los selectores de `getelementByidTag()` que toma un string en la entrada, retorna una cadena de string con todo los tag que ha filtrado.

No use búsqueda con clase tampoco con id, porque no hay manera para saber cuál será estos valores pero los tagname son todos iguales no son digitados por el programador.

Hay un ejercicio muy interesante la c, donde hay que seleccionar todos los párrafos primero y genera una lista de elementos después realizó otra búsqueda en esta lista para encontrar los recursos de imágenes.

Mi experiencia con esta práctica fue muy divertida:

Primero: realizó como simple es manipular el DOM y ya veo cómo será simple usar javascript para realizar este de tarea con los navegadores.

Segundo: casi usamos todas las herramientas para desarrollar aplicación web: manejo de paquete (gradle), intellij IDEA que me sorprende mucho que tan fácil es para trabajar con java y el uso de github para la entrega de la tarea fue una buena experiencia. Ahora solo falta trabajar con la base de datos.

Encuentro varias maneras para contestar una pregunta por ejemplo. La a) se puede hacer de transformar el documento en string con la función toString de los objeto de java y luego aplico el método split con los parámetro "\r\n|\r|\n" ya tengo un arreglo de string solo aplico arreglo.length para tener la cantidad de línea de recurso. La otra manera para hacer el mismo es usar el método stream luego aplicar el método count.

Espero que la siguiente práctica será mejor.

Bibliografia

<https://www.w3schools.com/>

<https://jsoup.org/cookbook/extracting-data/dom-navigation>