**PONTIFICIA UNIVERSIDAD CATÓLICA MADRE Y MAESTRA**

**FACULTAD DE CIENCIAS DE LAS INGENIERÍAS**

**DEPARTAMENTO DE INGENIERÍA EN SISTEMAS Y COMPUTACIÓN**



**Creación Cliente HTTP – Apache Components**

Materia: ISC-415.

*Profesor*: Carlos Camacho

Por

Ernesto Rodríguez

Matrícula:

2012-0201

Link de Git:

<https://github.com/er12/Prog-Web/tree/master/Practica1>

**Santiago de los Caballeros, República Dominicana**

**Mayo 2016**

# Introducción

El protocolo de transferencia de Hypertexto es el “medio de transporte” de archivos en la red de internet.  
El principal archivo utilizado en este protocolo es el tipo HTML (Hypert Text Markup Lenguage). En un principio era utilizado para crear referencias a otros archivos en una misma computadora, a través de links en las palabras que los identificaban.

Con la popularidad de esta forma comprartir información, nace los lenguajes JavaScript y CSS para adornar y manipular estos documentos.

En esta práctica se realizará un Cliente Http para analizar ciertas características de una página web.

# Desarrollo

El programa será escrito en Java bajo el constructor de proyectos Gradle.

El cliente se axiliará de la librería de Jsoup para extraer los elementos de la página que queramos.

Comenzamos por crear incluir las librerías a través de las dependencias de Gradle.  
La conexión comienza con:

doc = Jsoup.connect("http://www."+URL).get();

la cual extrae la información del url (localización de recursos uniformes/dirección de la página) y la convierte en un objeto Document.

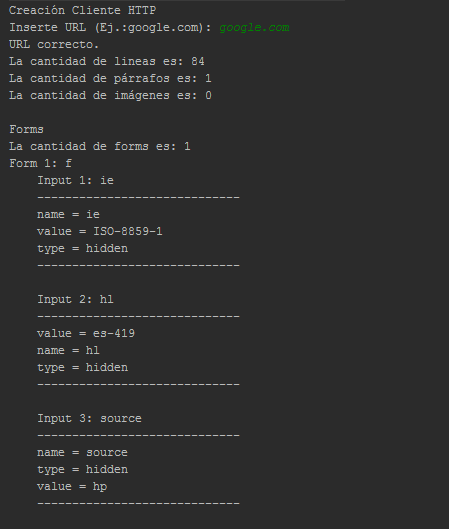
Un simple

Elements links = doc.getElementsByTag("p");

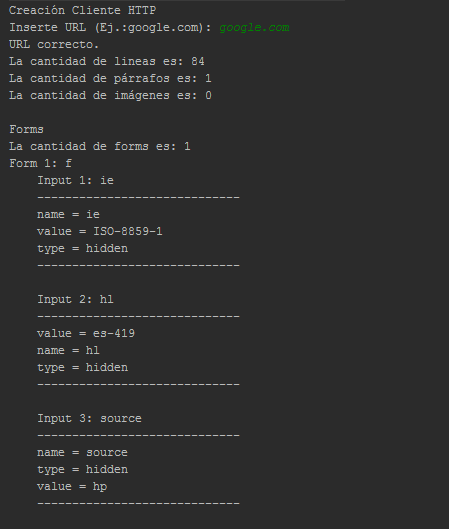
permite adquirir los tags (identificadores) del html que especifiquemos.

Para los atributos, Jsoup brinda una clase Attribute de los elementos que tiene la lista de atributos con keys (llaves) y valores.

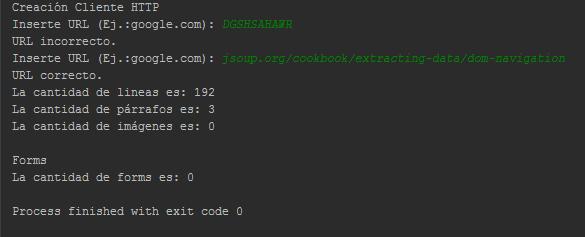
Esto es utilizado al tomar los atributos de los inputs pertenecientes a los forms.



El resultado es:



Si el URL es inválido, el programa no se conecta y, por ende, le pide al usuario escribir de nuevo.



# Conclusión

La librería Jsoup simplificó mucho la tarea de extraer la información deseada de la página de internet.

Se estudió con la herramienta de programador del navegador para comprobar los resultados.

La práctica terminó siendo sencilla y fácil de estudiar.

# Bibliografía

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