

WEBFINDER

OBTENCION DATOS HTTP



```
public class WebFinder {  ⤵ RodrigoMoreno2001 *  
  
    public static Web buscarWeb(String url,Integer puerto){ 1 usage  ⤵ RodrigoMoreno2001 *  
  
        Web web=null;  
  
        try {  
  
            HttpURLConnection conn = (HttpURLConnection) new URL( spec: url+": "+puerto).openConnection();  
  
            conn.setRequestMethod("GET");  
  
            conn.setRequestProperty("User-Agent", "Mozilla/5.0");  
  
            conn.setConnectTimeout(5000);  
            conn.setReadTimeout(5000);  
  
        }  
    }  
}
```

Para la obtención de datos http, usamos la Clase HttpURLConnection, usamos el metodo .setRequestMethod("GET") para obtener dichos datos.

OBTENCION DATOS HTTP



```
int codigoRespuesta = conn.getResponseCode();

String motor = conn.getHeaderField( name: "X-Powered-By");

if (motor == null) motor = "Desconocido";

String titulo = buscarTitulo(conn);

String servidor = conn.getHeaderField( name: "Server");

if (servidor == null) servidor = "Desconocido";

web=new Web(puerto,titulo,codigoRespuesta,motor,servidor);

} catch (Exception e) {
}
return web;
}
```

Una vez configurada la conexión hacia la web, esta clase nos permite obtener diferentes datos como el código de respuesta y algunos fragmentos del header

OBTENCION DATOS HTTP



```
public static String buscarTitulo(HttpURLConnection conn) throws IOException { 1 usage  🧑 RodrigoMoreno2001 *  
  
    String titulo;  
  
    try(BufferedReader reader = new BufferedReader(new InputStreamReader(conn.getInputStream()))){  
        String linea;  
        StringBuilder HTML = new StringBuilder();  
  
        while ((linea = reader.readLine()) != null) {  
            HTML.append(linea);  
            if(linea.contains("</title>")) break;  
        }  
        titulo=HTML.indexOf("<title>")==-1?"Desconocido":HTML.substring(HTML.indexOf("<title>")+7,HTML.indexOf("</title>"));  
    }  
    return titulo;  
}
```

Esta es una forma de obtener el titulo de una web, el método `.getInputStream()`, proporciona todo el HTML de la página, luego nosotros podemos interpretarlo

PREGUNTAS