

Instalación

<http://wiki.legido.com/doku.php?id=informatica:linux:selenium:instalacion>

Inspeccionar URL

Una página web tiene un DOM con diferentes zonas dependientes, como contenedores.

Cogeremos como ejemplo la web de instantbyte y buscaremos un artículo.

Vamos a ir a <http://www.instantbyte.com/> e inspeccionamos el código (F12 en firefox o botón derecho inspeccionar elemento) sobre la casilla de buscar

Abrimos navegador:

```
ipython3
```

```
from selenium import webdriver
from selenium.webdriver.firefox.options import Options
options = Options()

driver = webdriver.Firefox(options=options)

url = 'http://www.instantbyte.com/'
driver.get(url)
```

Formato xpath

Para seleccionar una zona usamos xpath, el formato es:

```
<span title="Armando Bronca">
```

```
xpath = '//*[@title="Armando Bronca"]'
o = driver.find_element_by_xpath(xpath)
```

Se puede modificar:

Contiene:

```
xpath = '//*[@contains(@title, "Armando Bronca")]'
```

Seleccionar zona

Nos ponemos en la web de instantbyte, en la casilla Búsqueda avanzada, Inspect Element

```
<body>
  <div class="cc_banner-wrapper">
  <table width="100%" cellpadding="0" cellspacing="0" border="0" align="center">
    <tbody>
      <tr>
        <td>
          <table style="width: 100%;>
            <table style="width: 100%;>
            <table style="width: 100%;>
          <table style="width: 100%;>
            <tbody>
              <tr>
                <td width="175" valign="top" align="left">
                  <!--start searchbox -->
                  <table class="infoBoxHeading" width="100%" cellpadding="0" cellspacing="0" border="0">
                    <tbody>
                      <tr>
                        <td>
                          <table class="infoBoxContents" width="100%" cellpadding="3" cellspacing="0" border="0">
                            <tbody>
                              <tr>
                                <td>
                                  <table class="boxText" align="center">
                                    <form name="quick_find" action="http://www.instantbyte.com/advanced_search_result.php" method="get">
                                      <input type="text" name="keywords" size="30" maxlength="30" style="width: 124px">
                                      <input type="hidden" name="osCsid" value="ttal824jldjmurcottm9j5lpb6">
                                      <br>
                                      <a href="http://www.instantbyte.com/advanced_search.php?osCsid=ttal824jldjmurcottm9j5lpb6">
                                    </form>
                                  </td>
                                </tr>
                              </tbody>
                            </table>
                          </td>
                        </tr>
                      </tbody>
                    </table>
                  </td>
                </tr>
              </tbody>
            </table>
          </td>
        </tr>
      </tbody>
    </table>
  </div>
```

Queremos seleccionar el campo input de tipo text. Aquí nos podemos encontrar con varios problemas si el dom está mal hecho como es el caso sin id, puede haber mas de uno:

```
xpath = '//*[@type="text"]'
o = driver.find_element_by_xpath(xpath)
```

Mostramos lo que ha seleccionado:

```
print(o.get_attribute('innerHTML'))
```

```
<input type="text" name="keywords" size="30" maxlength="30" style="width: 124px">
```

Ahora podemos enviar texto, hacer click.... Podemos borrar el contenido si suele tener algo:

```
o.clear()
```

Y enviamos lo que queramos

```
o.send_keys("omnitik")
```

Volvemos a borrar contenido:

```
o.clear()
```

Click

Abajo a la derecha tenemos el botón de “Lo tengo!” para las cookies. inspeccionamos:

```
<a href="#null" data-cc-event="click:dismiss" target="_blank" class="cc_btn cc_btn_accept_all">Lo tengo !</a>
```

Lo seleccionamos:

```
xpath = '//*[@data-cc-event="click:dismiss"]'
o = driver.find_element_by_xpath(xpath)

print(o.get_attribute('outerHTML'))
```

Hacemos click:

```
o.click()
```

Lista desplegable

Abajo a la izquierda donde pone Euro

```
<tr>
  <td class="boxText" align="center">
    <form name="currencies" action="http://www.instantbyte.com/index.php" method="get">
      <select name="currency" onchange="this.form.submit();" style="width: 100%"> ... </select>
      <input type="hidden" name="osCsid" value="ttal824jldjmursottm9j5lpb6">
    </form>
  </td>
</tr>
```

Si pulsamos los 3 puntos ... nos saldrán los posibles valores como option:

```
<tr>
  <td class="boxText" align="center">
    <form name="currencies" action="http://www.instantbyte.com/index.php" method="get">
      <select name="currency" onchange="this.form.submit();" style="width: 100%">
        <option value="ESP">Pesetas</option>
        <option value="EUR" selected="">Euro</option>
      </select>
      <input type="hidden" name="osCsid" value="ttal824jldjmursottm9j5lpb6">
    </form>
  </td>
</tr>
```

O si copiamos el código:

```
<select name="currency" onchange="this.form.submit();" style="width:
```

```
100%"><option value="ESP">Pesetas</option><option value="EUR"
selected="">Euro</option></select>
```

Vaya web he elegido que sale Pesetas.....

Para seleccionar una opción:

```
xpath = "//select[@name='currency']/option[@value='ESP']"
o = driver.find_element_by_xpath(xpath)
o.click()
```

Podemos seleccionar por texto, pero cuidado con las webs multidioma:

```
xpath = "//select[@name='currency']/option[text()='Pesetas']"
o = driver.find_element_by_xpath(xpath)
o.click()
```

Grabar todo el contenido de una web

```
content = driver.page_source
with open('webpage.html', 'w') as f:
    f.write(content)
```

From:

<http://wiki.legido.com/> - **Legido Wiki**

Permanent link:

<http://wiki.legido.com/doku.php?id=informatica:linux:selenium:ejemplos>

Last update: **2020/07/04 00:22**

