

Un método para obtener texto del internet y analizarlo

Selenium y las páginas ricas en javascript

Eric Magar

ITAM

1er Simposio TLH en las ciencias sociales
9/11/2017

Obtener/sistematizar historia legislativa de iniciativas en Chile

- 1 fecha de introducción
- 2 autor
- 3 comisión
- 4 informe
- 5 ...

Página web de la Cámara de Diputados de Chile

<http://camara.cl>

Obstáculo

javascript: información ausente del código, se “contruye” a través de solicitudes

Selenium

Módulo de Python que permite abrir un navegador, hacer login, emular el mouse y sus clics, etc.

Solución no es eficiente, pero es eficaz