

ACCESO INTELIGENTE A LA INFORMACIÓN

PRÁCTICA 4 – BEAUTIFULSOUP

Un uso que podemos darle a BeautifulSoup es el de crearnos una interfaz con la que automatizar las búsquedas de información y no tener que navegar una o varias webs. La idea es la de tener cargado en nuestro programa las URL de las páginas webs que nos interesen y mediante los parámetros de entrada hacer la búsqueda pertinente. Vamos a hacer un pequeño ejemplo.

Hay numerosas páginas con información de la agenda cultural y actividades en las ciudades, Por ejemplo en Sevilla la web:

<http://www.sevillaguia.com/sevillaguia/agendacultural/agendacultural.asp>.

Dentro de la misma hay dos partes:

Un índice a la izquierda clasificado por categorías

Una agenda cultural de noticias, con las fechas de celebración de los eventos.

Buscamos construir un programa con Tkinter con tres botones de opción:

- a) “Almacenar Categorías”, que usando BeautifulSoup que sea capaz de extraer todos los nombres y enlaces a los distintos elementos clasificados por categorías, y almacenarlos en una base de datos sqlite.
- b) “Buscar Categoría”, que muestre una ventana que permita al usuario introducir una categoría y muestre en otra ventana (en una listbox con scrollbar) los elementos y enlaces que hay en la misma, extrayéndolos de la BD.
- c) “Buscar Evento”, que muestre una ventana que permita al usuario introducir una palabra y muestre en otra ventana con las noticias de eventos que contienen dicha palabra.