

Práctica-Web Scraping

Ejercicio[10 puntos]

Considerar el ejemplo visto en la teoría para recuperar las fotos de trenes de un blog. En el ejemplo se han capturado las fotos que aparecían en la página principal del blog, sin embargo en la parte derecha del blog dispone en la parte derecha de las entradas del blog clasificadas por años.

Trenes y tiempos

Desde mediados de los 70 hasta finales de los 90, hice muchas fotos a trenes. En general, la calidad de las mismas no es mucha, pero todas tienen alguna pequeña historia detrás que seguro interesa a los amantes del ferrocarril. Voy a ir presentando esas fotos poco a poco y contaré, en lo que recuerde, su particular historia. A veces tendré que utilizar fotos de otros autores -a los que citaré y agradeceré- sobre todo cuanto más me adentre en "el vapor"

jueves, 12 de febrero de 2015

Su majestad el TER (RENFE 597-001/060) (II): La degradación a Regionales

Hacia 1985 varios acontecimientos empezaron a marcar el futuro de las 54 composiciones TER que estaban en servicio tras la baja por accidente de las 032, 035, 039, 041, 050 y 055. Se cerraron varias líneas deficitarias no electrificadas, los "camellos" 592 y 593 iban imponiéndose, la extraña relación entre plazas de segunda y primera clase restaba viajeros, la mecánica, y sobre todo la refrigeración y el sistema de frenado por zapatas, causaba cada vez mayores problemas y, además, estaba a punto de finalizar el contrato de mantenimiento con FIAT.



Archivo del blog

- ▼ 2015 (38)
- ▶ 09/27 - 10/04 (1)
- ▶ 09/20 - 09/27 (1)
- ▶ 09/13 - 09/20 (1)
- ▶ 09/06 - 09/13 (1)
- ▶ 08/23 - 08/30 (1)
- ▶ 08/09 - 08/16 (1)
- ▶ 07/26 - 08/02 (1)
- ▶ 07/19 - 07/26 (1)
- ▶ 07/12 - 07/19 (1)
- ▶ 07/05 - 07/12 (1)
- ▶ 06/28 - 07/05 (1)

Se pide:

- a) Crear un programa en Python que recorra todas las entradas del año 2015 (en total 38), cree un directorio en el disco local por cada entrada, y en cada directorio baje las imágenes de trenes de dicha entrada.
- b) Crear un minibuscador en Python que pregunte al usuario un conjunto de palabras clave, y recorra el blog buscando todas las entradas del blog dónde aparezcan esas palabras claves. Como resultado debe mostrar listadas las direcciones de las entradas dónde ha localizado esas palabras clave.

Normas de entrega

- Fecha tope de entrega: 10/11/2015
- La entrega se realizará subiendo un archivo zip con las soluciones realizadas en Python. El archivo tendrá como nombre WebScraping_GrupoX donde X será el número de grupo correspondiente.