

Ejercicio 3 Python 2.x

Objetivo: Manipulación de cadenas y uso de expresiones regulares.

Referencias:

Tutorial

<http://docs.python.org/tutorial/index.html>

String Processing:

<http://docs.python.org/library/string.html>

<http://docs.python.org/tutorial/introduction.html#strings>

Regular Expressions:

<http://docs.python.org/library/re.html>

<http://docs.python.org/howto/regex.html#regex-howto>

Construir un tokenizador para el español, que, dado un fichero de texto de entrada (*entrada_tokenizador.txt*), separe en tokens, y lo muestre en un fichero de salida en el formato que se muestra en (*salida_tokenizador.txt*). Se deben cumplir las siguientes restricciones:

- 1) Los símbolos que hay que separar de cada palabra son: () . , ' " ? ¿ ! | ... ; :
- 2) No se deben separar los números decimales, ejemplo: 44,45 45.60
- 3) No se deben separar fechas 12/12/90, 12-03-98, ni horas, 9:30 h.
- 4) No se deben separar direcciones web <http://www.colorin.com> ni de correo electrónico xx@cdit.com