



## Trabajo Práctico Introducción a la Programación de Protocolos de la Pila TCP/IP con Python

### Bibliografía (ampliada)

El tutorial de Python. <http://docs.python.org.ar/tutorial/pdfs/TutorialPython2.pdf>

Python Standard Library. Fredrik Lundh. O'Reilly & Associates. <http://effbot.org/zone/librarybook-ora.htm>

Documentación Módulos de Internet: <http://docs.python.org/library/internet.html>

**Fecha de Entrega:** 14/09/2015

**Objetivo:** Conocer los fundamentos de la programación en lenguaje Python y experimentar con las facilidades para el manejo de protocolos de la pila TCP/IP en alto nivel (utilizando módulos del lenguaje).

1. Codifique un script que realice un búsqueda en Wikipedia utilizando `urllib`, el programa debe ser llamado por línea de comandos y no debe ser interactivo todos los parámetros para la búsqueda y configuración deben ser tomados de `argv` o en su defecto utilizando el módulo `ConfigParser`. El resultado debe ser guardado en un archivo de texto.
2. Escriba una segunda rutina para el ejercicio 1 donde, se verifique en la página descargada si hay sugerencia de desambiguación y en el caso de existir la misma recupere la página obtenga las sugerencias y descargue cada una de ellas en archivos separados.
3. Codifique un script que tome una cadena como parámetro, ejecute una búsqueda en Yahoo, ASK y en Google. Investigue qué opciones debe utilizar en cada caso. Comentar brevemente qué módulos utilizó y con qué dificultades se ha encontrado. Utilice además para la llamada el comando `time` para controlar el tiempo de ejecución de su proceso.
4. Codifique un script que reciba como parámetro un archivo con canales de noticias (RSS) y descargue y parsee el resultado descargando el contenido multimedia y en el caso de `iframe` de gadgets de videos también los descargue. Luego, genere una página html que tenga las noticias ordenadas por fecha de todos los feeds y el contenido multimedia en el caso que tenga. Explique detalladamente el procedimiento realizado, cuál es la estructura de almacenamiento de todos los contenidos multimedia y los textos. **Aclaración: No se puede utilizar librerías que lean canales RSS.**
5. Investigue el módulo `BaseHTTPServer` y escriba un programa que permita recuperar una página html básica escrita por ud. El servidor http debe generar un log con información relevante de cada petición y volcarla a un archivo a tal efecto.
6. Escriba un programa que permita sincronizar dos directorios remotos vía protocolo HTTP. Explique someramente cómo trabaja su aplicación, sus capacidades y sus limitaciones.
7. Pensemos que estamos gestionando un servidor Web que atiende a un gran número de usuarios, en estos casos es posible mejorar el rendimiento de nuestro servicio utilizando balanceo de carga. Investigue esta técnica y explique someramente en qué consiste y con qué algoritmos puede ser atacada. Luego, escriba un programa que trabaje como Load Balancer de al menos 3 nodos http internos. Nuestro balanceador debe soportar al menos un algoritmo de Round Robin.