

SOAP desde Python

SOAP al igual que XML-RPC puede ser implementado en la mayoría de los lenguajes de programación actuales dentro del entorno Linux. Entre ellos, está Python, lenguaje de script similar a Perl, muy querido entre los Debianitas.

Utilizando un módulo Python, podemos utilizar toda la potencia de los Servicios Web basados en SOAP. Veamos como:

```
wget http://switch.dl.sourceforge.net/sourceforge/pywebsvcs/SOAPpy097.tgz
tar -zxvf SOAPpy097.tgz
cp -Rdf SOAPpy097 /usr/lib/python2.2/site-packages
```

Es recomendable utilizar Python 2.1, o 2.2, para evitar un problema de memoria que se da con el parser XML de la versión 2.0

Referencia: <http://www.fettig.net/googleRank.txt>

Un script en Python que utilice este módulo para efectuar una consulta sobre Google sería así:

```
#!/usr/bin/python
import SOAP
import string

key = 'iwnUXUtHj3bteg5FWfBJDwui3SPeB+iy'
soapfalse = SOAP.booleanType(0)
emptyStr = SOAP.stringType("")

def rank(site, query):
    x=1
    server = SOAP.SOAPProxy('http://api.google.com/search/beta2', namespace=
'urn:GoogleSearch')
    start = 0
    while start < 10:
        data = server.doGoogleSearch( SOAP.stringType(key),
SOAP.stringType(query), start*10, 10, soapfalse, emptyStr, soapfalse, emptyStr,
emptyStr, emptyStr)
        results = data.resultElements
        for result in results:
            try:
                result.URL.index(site)
                return (result.URL, start*10 + results.index(result) + 1)
            except:
                pass
            x = x+1

        start = start + 1

if __name__ == '__main__':
    import sys
    searchPhrase = string.join(sys.argv[2:], " ")
    result = rank(sys.argv[1], searchPhrase)
    if result:
        print "" + result[0] + "" + " is currently ranked #" + str(result[1])
+ " on Google for the search phrase " + searchPhrase + "."
    else:
        print "" + sys.argv[1] + "" + " is not currently in the top 100
results for the search phrase " + searchPhrase + " on Google."
```

Una versión más simplificada y más clara de este código es la siguiente:

```
#!/usr/bin/python
import SOAP
key = 'iwnUXUtHj3bteg5FWfBJDwui3SPeB+iy'
query = 'python'
s = SOAP.SOAPProxy("http://api.google.com/search/beta2")
r = s.doGoogleSearch(key, query, 0, 10, "false", "", "false", "", "latin1",
"latin1")
```

Existe un proyecto llamado [PyGoogle](#) que puede ayudarnos a la hora de extraer información adicional de Google, como por ejemplo páginas cacheadas o sugerencias de corrección ortográfica. PyGoogle es una aplicación basada en SOAP y el Servicio Web de Google.