

Distributed Software Development
Using Google with SOAP

Chris Brooks

Department of Computer Science
University of San Francisco

9-2: Using Google's Web Services

- Google also provides a Web Services interface to their data.
- They use SOAP, rather than REST.
 - It's possible to send a URL from a program, but results are returned in HTML, rather than XML.
- This was a big point of contention when Google put this out - they'd previously had a REST-style interface, which is no longer freely available.

9-3: SOAP review

- Recall that SOAP uses a procedural model, rather than a data-retrieval model.
- Method/procedure names, input and output parameters are encoded in the body of a SOAP message.
- Service API is exposed as a WSDL file
 - A particular flavor of XML Schema.
- This allows clients to automatically generate code to interface with the service.

9-4: Interfaces

- The Google API toolkit has examples of using their service with Java and C#.
- Dive Into Python shows how to use it with Python.
- Ruby/Google seems to work, but I haven't used it.

9-5: The Google WSDL

- There are three client-side procedures that can be used:
 - doGoogleSearch
 - doSpellingSuggestion
 - doGetCachedPage
- doGoogleSearch is probably the most commonly-used one.

9-6: Using the Google API with Python

- Here's a simple example, adapted from Dive Into Python

```
#!/usr/bin/python
from SOAPpy import WSDL
mykey = 12345 # get your own key
search = WSDL.Proxy('path/to/google.wsdl')

results = search.doGoogleSearch(key, buffy, 0, 10, False, "", False,
"", "utf-8", "utf-8")
for result in results.resultElements :
    print result.title, result.URL, result.snippet
```

9-7: Inputs

- doGoogleSearch takes nine arguments:
 - Your API key
 - Your search string
 - The starting result number
 - The ending result number
 - filter - remove duplicates?
 - restrict- you can set this to country+ to get results from particular countries, e.g. 'countryUK'
 - safeSearch
 - lr - choose result language
 - ie, oe - encoding of input/output - should be 'utf-8'

9-8: Outputs

- This returns a dictionary-like object
 - searchTips
 - endIndex
 - searchQuery
 - resultElements
 - estimatedTotalResultsCount
 - searchTime

9-9: Result Elements

- ResultElements is a list of results
- Each result has:
 - cachedSize
 - title
 - URL
 - snippet
 - summary
 - directoryCategory

9-10: Using the Google Web Service with Java

- The calls and results look very similar - take a look at the javadoc for examples.
- (Ruby folks - the calls are very similar. Ruby/Google is not incredibly well-documented, but you should be used to that by now :-)

9-11: Summary

- Much of the complexity of SOAP can be (usually) hidden by tools that automatically generate proxy code from a WSDL.
- This gives the user a programmatic interface like the one they'd get for working with any other library.
- Google's API is pretty simple and straightforward to use.