

Web-based APIs

hussein suleman
uct cs honours 2006

Google Web Search API

- Google currently has a free beta Web Service interface.
 - Up to 1000 requests per day per user.
 - Defined using WSDL, using HTTP-SOAP.
- Every user must first request a unique key to track user activity, and provide this key with every request.
- Requests can be for search results, cached data or spelling checks.

Google WSDL 1/2

```
<xsd:complexType name="GoogleSearchResult">
  <xsd:all>
    <xsd:element name="documentFiltering" type="xsd:boolean"/>
    <xsd:element name="searchComments" type="xsd:string"/>
    <xsd:element name="estimatedTotalResultsCount" type="xsd:int"/>
    <xsd:element name="estimateIsExact" type="xsd:boolean"/>
    <xsd:element name="resultElements" type="typens:ResultEl"/>
    <xsd:element name="searchQuery" type="xsd:string"/>
    <xsd:element name="startIndex" type="xsd:int"/>
    <xsd:element name="endIndex" type="xsd:int"/>
    <xsd:element name="searchTips" type="xsd:string"/>
    <xsd:element name="directoryCategories" type="typens:Director"/>
    <xsd:element name="searchTime" type="xsd:double"/>
  </xsd:all>
</xsd:complexType>

<xsd:complexType name="ResultElement">
  <xsd:all>
    <xsd:element name="summary" type="xsd:string"/>
    <xsd:element name="URL" type="xsd:string"/>
    <xsd:element name="snippet" type="xsd:string"/>
    <xsd:element name="title" type="xsd:string"/>
    <xsd:element name="cachedSize" type="xsd:string"/>
    <xsd:element name="relatedInformationPresent" type="xsd:boolean"/>
    <xsd:element name="hostName" type="xsd:string"/>
    <xsd:element name="directoryCategory" type="typens:DirectoryCategory"/>
    <xsd:element name="directoryTitle" type="xsd:string"/>
  </xsd:all>
</xsd:complexType>
```

Google WSDL 2/2

```
<message name="doGoogleSearch">
  <part name="key" type="xsd:string"/>
  <part name="q" type="xsd:string"/>
  <part name="start" type="xsd:int"/>
  <part name="maxResults" type="xsd:int"/>
  <part name="filter" type="xsd:boolean"/>
  <part name="restrict" type="xsd:string"/>
  <part name="safeSearch" type="xsd:boolean"/>
  <part name="lr" type="xsd:string"/>
  <part name="ie" type="xsd:string"/>
  <part name="oe" type="xsd:string"/>
</message>

<operation name="doGoogleSearch">
  <input message="typens:doGoogleSearch"/>
  <output message="typens:doGoogleSearchResponse"/>
</operation>
```

Google SOAP Request

```
<SOAP-ENV:Envelope xmlns:SOAP-
  ENV="http://schemas.xmlsoap.org/soap/envelope/" xm
  lns:xsi="http://www.w3.org/1999/XMLSchema-instance"
  xmlns:xsd="http://www.w3.org
  /1999/XMLSchema">
  <SOAP-ENV:Body>
    <ns1:doGoogleSearch xmlns:ns1="urn:GoogleSearch"
      SOAP-
      ENV:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/">
      <key
        xsi:type="xsd:string">00000000000000000000000000000000</key>
      <q xsi:type="xsd:string">uct computer science</q>
      <start xsi:type="xsd:int">0</start>
      <maxResults xsi:type="xsd:int">10</maxResults>
      <filter xsi:type="xsd:boolean">>false</filter>
      <restrict xsi:type="xsd:string"></restrict>
      <safeSearch xsi:type="xsd:boolean">>false</safeSearch>
      <lr xsi:type="xsd:string"></lr>
      <ie xsi:type="xsd:string">latin1</ie>
      <oe xsi:type="xsd:string">latin1</oe>
    </ns1:doGoogleSearch>
  </SOAP-ENV:Body>
</SOAP-ENV:Envelope>
```

Google SOAP Response 1/2

```
<SOAP-ENV:Envelope>
  <SOAP-ENV:Body>
    <ns1:doGoogleSearchResponse>
      <return xsi:type="ns1:GoogleSearchResult">
        <documentFiltering xsi:type="xsd:boolean">>false</documentFiltering>
        <estimatedTotalResultsCount
          xsi:type="xsd:int">3</estimatedTotalResultsC
          ount>
        <directoryCategories
          xmlns:ns2="http://schemas.xmlsoap.org/soap/encoding
          /" xsi:type="ns2:Array"
          ns2:arrayType="ns1:DirectoryCategory[0]"></directoryCate
          gories>
        <searchTime xsi:type="xsd:double">0.194871</searchTime>
        <resultElements
          xmlns:ns3="http://schemas.xmlsoap.org/soap/encoding/" xs
          i:type="ns3:Array" ns3:arrayType="ns1:ResultElement[3]">
          <item xsi:type="ns1:ResultElement">
            <cachedSize xsi:type="xsd:string">12k</cachedSize>
            <hostName xsi:type="xsd:string"></hostName>
            <snippet xsi:type="xsd:string"> some stuff ... </snippet>
            <directoryCategory xsi:type="ns1:DirectoryCategory">
              <specialEncoding xsi:type="xsd:string"></specialEncoding>
              <fullViewableName xsi:type="xsd:string"></fullViewableName>
            </directoryCategory>
          </item>
        </resultElements>
      </return>
    </ns1:doGoogleSearchResponse>
  </SOAP-ENV:Body>
</SOAP-ENV:Envelope>
```

Google SOAP Response 2/2

```
<relatedInformationPresent xsi:type="xsd:boolean">true</relatedInformationPresent>
  <directoryTitle xsi:type="xsd:string"></directoryTitle>
  <summary xsi:type="xsd:string"></summary>
  <URL
    xsi:type="xsd:string">http://hci.stanford.edu/cs147/examples/shrdlu/</URL>
  <title xsi:type="xsd:string">&lt;b&gt;SHRDLU&lt;/b&gt;</title>
</item>
</resultElements>
<endIndex xsi:type="xsd:int">3</endIndex>
<searchTips xsi:type="xsd:string"></searchTips>
<searchComments xsi:type="xsd:string"></searchComments>
<startIndex xsi:type="xsd:int">1</startIndex>
<estimateIsExact xsi:type="xsd:boolean">true</estimateIsExact>
<searchQuery xsi:type="xsd:string">shrdlu winograd maclisp
teletype</searchQuery>
</return>
</nsl:doGoogleSearchResponse>
</SOAP-ENV:Body>
</SOAP-ENV:Envelope>
```

Using Google API from Perl

```
use SOAP::Lite;
# create soap object
my $key='00000000000000000000000000000000';
my $query="uct computer science";
my $googleSearch = SOAP::Lite -> service("file:./GoogleSearch.wsdl")
    -> proxy
    ('http://api.google.com/search/beta2',
     proxy => ['http' =>
       'http://cache.uct.ac.za:8080']);
# submit to server and get results / retry while request not accepted
my $result;
my $max_retries = 0;
my $retry_count = 0;
while (!$result->{resultElements} && $retry_count <= $max_retries)
{
    eval {
        $result = $googleSearch -> doGoogleSearch(
            $key, $query, 0, 10, "false", "", "false", "", "latin1", "latin1");
    };
    $retry_count++;
}
# print out titles and URLs
foreach my $aresult (@{$result->{resultElements}})
{
    print (sprintf ("%s\n    URL: %s\n", $aresult->{title'}, $aresult->{URL'}));
}
}
```

Output from Google via Perl

Department of **Computer Science**, University of Cape Town
URL: <http://www.cs.uct.ac.za/>
Why study at the Department of **Computer Science**, University of **...**
URL: <http://www.cs.uct.ac.za/newstudents/why.html>
Prospective Undergrad Students for the Department of **Computer** **...**
URL: <http://www.cs.uct.ac.za/newstudents/staffp.html>
About Us: Department of **Computer Science**, University of Cape Town
URL: <http://www.cs.uct.ac.za/about.html>
Postgrad Students of the Department of **Computer Science** **...**
URL: <http://www.cs.uct.ac.za/people/stud-postgrad.html>
Prospective Undergrad Students for the Department of **Computer** **...**
URL: <http://www.cs.uct.ac.za/people/stud-alumni.html>
Admin staff of the Department of **Computer Science**, University of **...**
URL: <http://www.cs.uct.ac.za/people/staff-admin.html>
Department of **Computer Science**
URL: <http://moodle.cs.uct.ac.za/>
Using Windows CE.NET 4.1 in **UCT's** **Computer Science** Department
URL:
http://arc.cs.odu.edu:8080/dp9/getrecord/oai_dc/techreports.cs.uct.ac.za/oai:techreports.cs.uct.ac.za:65
ScienceDirect - Theoretical **Computer Science** : Scheduling UET-**UCT**
...
URL: [http://dx.doi.org/10.1016/0304-3975\(96\)00035-7](http://dx.doi.org/10.1016/0304-3975(96)00035-7)

Beyond Google Web Search

- ❑ Other Google Web APIs
 - Google Maps
 - Google Earth
 - etc.
- ❑ Google RSS/Atom feeds for GMail.

- ❑ MSN and Yahoo Web APIs.

Amazon Electronic Commerce

- A full suite of Web Services to interface with Amazon.com
 - Both SOAP and REST (URL-encoded) interfaces
 - Many more interfaces than Google. Why?
 - Fewer restrictions on usage, in production and free for many purposes!

Amazon REST Request

- `http://webservices.amazon.com/onca/xml?Service=AWSECommerceService&AWSAccessKeyId=0000000000&Operation=ItemLookup&ItemId=0620056886`
- baseURL
 - `http://webservices.amazon.com/onca/xml?Service=AWSECommerceService`
- Key to access service
 - `AWSAccessKeyId=000000000000`
- Operation to perform
 - `Operation=ItemLookup`
- Item to display
 - `ItemId=0620056886`

Output from Amazon 1/2

```
<ItemLookupResponse>
  <OperationRequest>
    <HTTPHeaders>
      <Header Name="UserAgent" Value="Mozilla/5.0 (Windows; U; Windows
        NT 5.1; en-US; rv:1.8.0.5) Gecko/20060719 Firefox/1.5.0.5"/>
    </HTTPHeaders>
    <RequestId>0XJBJRVTVME4XENAVPQS</RequestId>
    <Arguments>
      <Argument Name="Service" Value="AWSECommerceService"/>
      <Argument Name="ItemId" Value="0620056886"/>
      <Argument Name="AWSAccessKeyId" Value="00000000000000000000"/>
      <Argument Name="Operation" Value="ItemLookup"/>
    </Arguments>
    <RequestProcessingTime>0.0196020603179932</RequestProcessingTime>
  </OperationRequest>
  <Items>
    <Request>
      <IsValid>True</IsValid>
      <ItemLookupRequest>
        <ItemId>0620056886</ItemId>
      </ItemLookupRequest>
    </Request>
  </Items>
</ItemLookupResponse>
```

Output from Amazon 2/2

```
<Item>
  <ASIN>0620056886</ASIN>
  <DetailPageURL>
    http://www.amazon.com/exec/obidos/redirect?tag=ws%26link_code=xm2%26camp=20
    25%26creative=165953%26path=http://www.amazon.com/gp/redirect.html%253fA
    SIN=0620056886%2526tag=ws%2526lcode=xm2%2526cID=2025%2526ccmID=165953%25
    26location=/o/ASIN/0620056886%25253FSubscriptionId=00000000000000000000
  </DetailPageURL>
  <ItemAttributes>
    <Author>Andrew Verster</Author>
    <Creator Role="Editor">Zuleikha Mayat</Creator>
    <Creator Role="Photographer">Dennis Bughwan</Creator>
    <Creator Role="Illustrator">Nalin Bughwan</Creator>
    <Manufacturer>Domain Enterprises</Manufacturer>
    <ProductGroup>Book</ProductGroup>
    <Title>Indian Delights</Title>
  </ItemAttributes>
</Item>
</Items>
</ItemLookupResponse>
```

References

- Amazon.com (2006) Looking up Product Data in the Amazon Catalog. Website <http://docs.amazonwebservices.com/AWSEcommerceService/2006-06-28/>
- Google (2006) Google SOAP Search API Reference. Website <http://www.google.com/apis/reference.html>