

Google Web API

- Einführung
- Wichtige Grundlagen
- Funktionsweise
- Möglichkeiten und Beschränkungen

- WSDL im Detail und Praktische Beispiele (2. Teil, Vortrag von Rafael)



Die GoogleAPI [1]

- Google Web API ist ein Web Service von Google
 - Stellt eine Schnittstelle zum Google-Index bereit
- Dient (noch) zum Automatisieren von Suchanfragen und deren Auswertung
- Vor 2002:
 - Grundsätzlich keine automatischen Anfragen an Google erlaubt
 - IP-Adressen wurden gezielt geblockt
- Nach 2002:
 - Wurde **offiziell** bereitgestellt
 - Automatische Anfragen und Antworten erlaubt



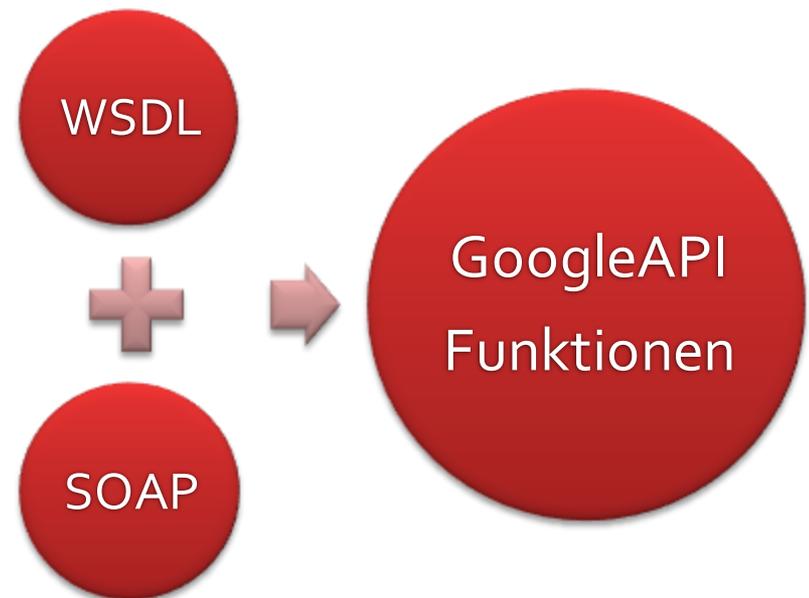
Die GoogleAPI [2]

- Google Web API als Ersatz von selbst erstellten Programmen
 - u.U. sind diese mächtiger als der Service von Google
- Trotz der Bereitstellung wird der Google Index nicht zu 100% frei zugänglich für die automatische Suchanfrage gemacht
 - Beschränkt auf wenige Funktionen
- Verhinderung von eigenen Programmen nicht ganz möglich
 - Aber:** Durch Vergabe von „Developer Keys“ ⇒
 - Identifizierung der Suchanfragen möglich bzw.
 - Zumindest kann man Sie zuordnen



Grundlagen

- Um mit der GoogleAPI arbeiten zu können grundsätzlich Informationen nötig:
 1. Welche Methoden, Funktionen etc. stehen zur Verfügung
 2. Wie kann man diese nutzen
- WSDL [whiz-dill]
 - Beschreibt was zur Verfügung steht
- SOAP
 - Ruft die eigentlichen Funktionen auf



WSDL / SOAP

WSDL (Web Services Description Language)

- Definiert eine XML-Spezifikation
- Beschreibt Web Services
 - Operationen, Parameter und Rückgabewerte

SOAP (Simple Object Access Protocol)

- Netzwerkprotokoll (HTTP, TCP)
- Nachrichten werden auf Basis von XML ausgetauscht

WSDL definierte Funktionen

- Eine XML-Datei mit integriertem XML-Schema

```
<portType name="GoogleSearchPort">  
  
  <operation name="doGetCachedPage">  
    <input message="typens:doGetCachedPage" />  
    <output message="typens:doGetCachedPageResponse" />  
  </operation>  
  
  <operation name="doSpellingSuggestion">  
    <input message="typens:doSpellingSuggestion" />  
    <output  
message="typens:doSpellingSuggestionResponse" />  
  </operation>  
  
  <operation name="doGoogleSearch">  
    <input message="typens:doGoogleSearch" />  
    <output message="typens:doGoogleSearchResponse" />  
  </operation>  
  
</portType>
```

Abstrakte
Definition

SOAP Nachrichten

- SOAP Nachrichtenformat (Requests & Responses)

```
<?xml version="1.0"?>
<s:Envelope xmlns:s="http://www.w3.org/2001/12/soap-envelope">
  <s:Body>
    <m:TitleInDatabase xmlns:m="http://www.seminar-db.de/soap">
      DOM, SAX und SOAP
    </m:TitleInDatabase>
  </s:Body>
</s:Envelope>
```



SOAP Nachrichten

- SOAP Nachrichtenformat (Requests & Responses)

```
<?xml version="1.0" ?>
<
  <?xml version="1.0" ?>
  <s:Envelope xmlns:s="http://www.w3.org/2001/12/soap-envelope">
    <s:Header>
      <m:RequestID xmlns:m="http://www.lecture-
db.de/soap">a3f5c109b</m:RequestID>
    </s:Header>
    <s:Body>
      <m:DbResponse xmlns:m="http://www.lecture-db.de/soap">
        <m:title value="DOM, SAX und SOAP">
          <m:Choice value="1">Arbeitsbericht Informatik</m:Choice>
          <m:Choice value="2">Seminar XML und Datenbanken</m:Choice>
        </m:title>
      </m:DbResponse>
    </s:Body>
  </s:Envelope>
```

Möglichkeiten

- Konkret bietet Google API folgende Funktionen:
 - Einfache Suche
 - Rechtschreibung bzw. richtige Schreibweise vorschlagen lassen
 - Ergebnisse vom Google-Cache anfordern
- Was ist noch zu beachten?
 - Mehr als 10 Ergebnisse anzeigen lassen ohne programmiertechnische Tricks nicht möglich
 - Pro Tag „nur“ 1000 Anfragen erlaubt



Beschränkungen

- Wichtig zu wissen
 - Die Google API wurde damals als Beta vorgestellt
 - Seit Anfang 2007 findet kein Support von Google mehr statt
- Neue Schlüssel werden auch nicht mehr vergeben
 - Glücklicherweise haben wir noch einen alten Schlüssel 
 - Bis auf weiteres funktionieren noch die alten Schlüssel
- Alternativen
 - Yahoo Search API (Developer Key vorhanden)
 - Eigenes Script (normale Google Suchantwort muss auseinander genommen werden)



Vielen Dank
und weiter geht's...

