



GazetteerWeb

Interface zur Abfrage eines
Geo-Webservice

Jens Angerstein
Karl-Heinz Deutinger
Frank Kleine
Danny Körnig
Jörg Schönebaum



Agenda

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- Einleitung
- Funktion eines Gazetteer
- Schnittstelle
- Cache
- Speicherung
- Merge von Abfrageergebnissen
- insertNodePhrase
- Fazit & Ausblick



Einleitung

Aufgabenstellung

- Abfragen eines Gazetteer-Webservices mittels PHP-Klasse
- Caching der Anfragen
- Bereitstellen von Verknüpfungsfunktionen für die Gazetteer-Ergebnisse

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick



Funktion eines Gazetteer

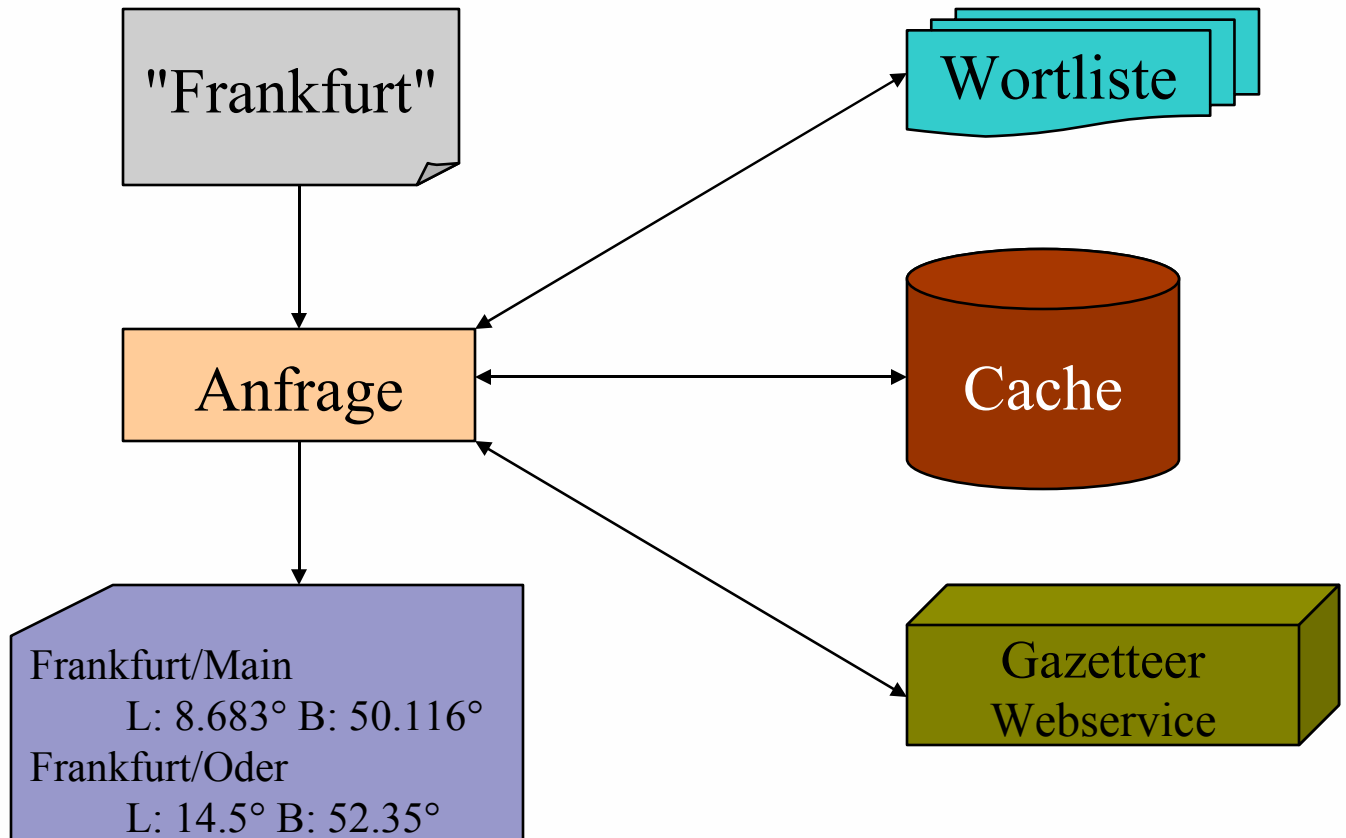
- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- liefert geographische Merkmale (Breiten-/Längengrad, ...)
- Alexandria Digital Library Project
 - XML-Anfragesprache
 - Längen- und Breitengrad
 - liegt-In-Beziehung
 - historische Namen
 - ...



Schnittstelle

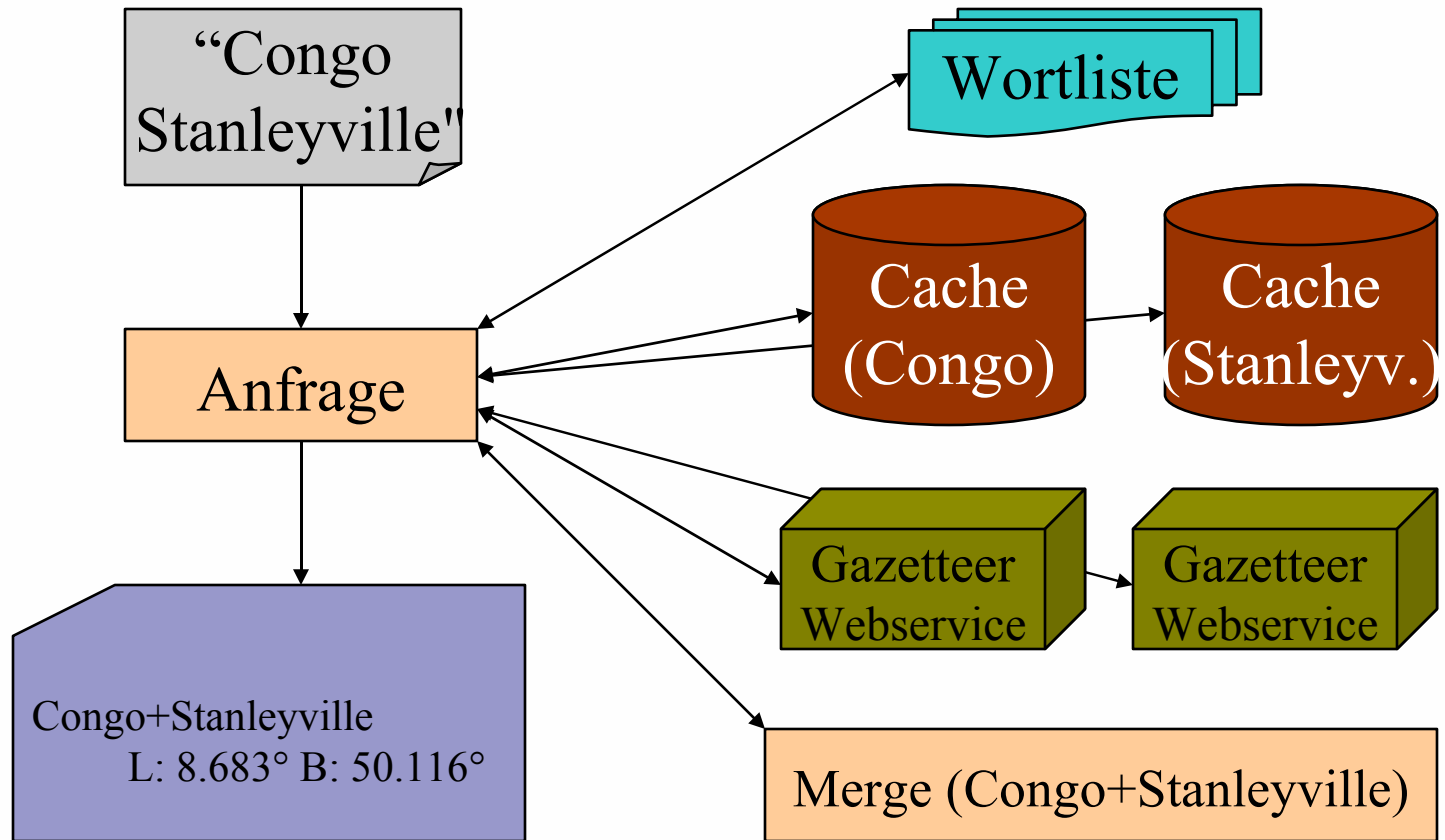
einfacher Ablauf





Schnittstelle

komplexer Ablauf



- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick



Cache

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- Vier Konfigurationsmöglichkeiten:
 - Maximale Lebenszeit
 - Minimale Lebenszeit
 - Minimale Nutzungszahl
 - Maximale Cache-Größe
- Speicherung als serialisiertes Objekt statt des XML-Abfrageergebnisses
 - » günstiger für Performance

```
serialize($result);
```



Speicherung

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- Ergebnisspeicherung in temporären Tabellen: Identifier, Name, Koordinaten, Relevanz
- Relevanz wird durch GazetteerWeb eingeführt: alle gelieferten Ergebnisse haben Relevanz = 1
- Verwendung dieser Tabellen für Ergebnisverarbeitung (`insertNodePhrase()`, Merge)



Merge von Abfrageergebnissen

■ ellenlange SQL-Abfragen

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick



insertNodePhrase()

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- Speichern der Dokumentnamen und Knoten der Geoinformationen
- Vorteile:
 - Schnelles Auffinden von bestimmten Dokumenten
 - Geo-Index über die Dokumente
- Nachteil
 - Redundante Speicherung der Daten im Cache und im Geo-Index



Fazit

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- GazetteerWeb lauffähiges Programm zur Unterstützung der Suche nach Geo-Informationen
- Genaues Klassifizieren der Geo-Informationen möglich
- Effiziente Anfragen durch Caching-System
- Gute Bedienbarkeit über ein Interface



Ausblick

- Einleitung
- Gazetteer?
- Schnittstelle
- Cache
- Speicherung
- Merge
- insertNodePhrase'
- Fazit & Ausblick

- Anbindung weiterer Gazetteer (XML)
- Erweitertes Relevanz-Ranking
- Einsatz als zusätzlicher Webservice (Visualisierung)
- Einsatz als Proxy zur Reduzierung der Gazetteeranfragen



Vorführung

Und so funktioniert es praktisch...

Magdeburg

Query

GAZ_TEMPORAER_MAGDEBURG						
ID	Name	X1	Y1	X2	Y2	Relevanz
adlgaz-1-2169188-5a	Magdeburg-Sudost, Stadtbezirk	11.650000	52.100000	11.950000	52.600000	1.00
adlgaz-1-2218715-5a	Sudenburg	11.600000	52.100000	11.600000	52.100000	1.00
adlgaz-1-2102065-61	Buckau	11.633330	52.083330	11.633330	52.083330	1.00
adlgaz-1-2117153-0a	Fermersleben	11.650000	52.083330	11.650000	52.083330	1.00
adlgaz-1-2179808-50	Neustadt	11.633330	52.150000	11.633330	52.150000	1.00
adlgaz-1-2191644-63	Prester	11.666670	52.100000	11.666670	52.100000	1.00
adlgaz-1-2199844-64	Rothensee	11.666670	52.183330	11.666670	52.183330	1.00
adlgaz-1-2107027-03	Diesdorf	11.566670	52.133330	11.566670	52.133330	1.00
adlgaz-1-2187811-64	Ottersleben	11.583330	52.100000	11.583330	52.100000	1.00
adlgaz-1-2103179-48	Burg	11.850000	52.266670	11.850000	52.266670	1.00
adlgaz-1-2169193-19	Magdeburg	11.666670	52.166670	11.666670	52.166670	1.00
adlgaz-1-2169187-4a	Magdeburg, Stadtkreis	9.000000	53.000000	9.000000	53.000000	0.50

im merge benutzen Mergel [Tabelle löschen](#)

GAZ_TEMPORAER_MAGDEBURG1						
ID	Name	X1	Y1	X2	Y2	Relevanz
adlgaz-1-2169193-19	Magdeburg	11.666670	52.166670	11.666670	52.166670	1.00

im merge benutzen Mergel [Tabelle löschen](#)