

Übung Netzbasierte Informationssysteme

Termin 10: Semantisches Geoinformationssystem



Prof. Dr. Adrian Paschke

Arbeitsgruppe Corporate Semantic Web (AG-CSW)
Institut für Informatik, Freie Universität Berlin
paschke@inf.fu-berlin.de
<http://www.inf.fu-berlin.de/groups/ag-csw/>

- **Lernziele des 9. Übungsblattes waren:**
 - Kennenlernen des Engineerings von Ontologien
- **Wie war`s?**
 - Fehlerquellen
 - Probleme?

Übung 10: Semantische Geoinformationssystem



Read, Laughed and Filed: Why Siemens Should Never Have Been Listed by Wall Street

Germany's scandal-plagued Siemens should have been barred from being listed on the US stock exchange in 2001. According to internal reports, the company's management at the time, under then-CEO Heinrich von Plerer, was guilty of serious negligence.

SPIEGEL ONLINE SPIEGEL ONLINE - International Madison, United States

The screenshot shows a satellite map of North America with several red 'SPIEGEL ONLINE' logos overlaid on the continent. A small inset map in the top left shows the location of the article on a world map. A north arrow is visible in the top right corner.



Madagascar, 27 days ago - Several New Species in Ensembl

Ensembl now contains data on the Rabbit , Small Madagascar Hedgehog, Nine-Banded Armadillo and Yellow Fever Mosquito ... more. {www.ebi.ac.uk}

Several New Species in Ensembl
Madagascar, 27 days ago

The screenshot shows a satellite map of Madagascar with a yellow dot on the island. A text box is overlaid on the map, containing the title and a list of species. A link to the Ensembl website is provided.

- Extraktion von Ereignis- (Konferenzen/Workshops) und Publikationsinformationen aus Webdatenbanken
- Bestimmung der Geoinformationen (Veranstaltungsort, Ort der Erscheinung)
- Visualisierung der extrahierten Webinformationen auf einer interaktiven 3D Weltkugel am korrekten Ort und in Realzeit
- (Konfigurationsmöglichkeiten des Semantischen Geoinformationssystems für den Nutzer zur Beschreibung des Informationsbedarfes – Personal Information Manager)
- Hinweis: Verwendung bestehender Tools / APIs

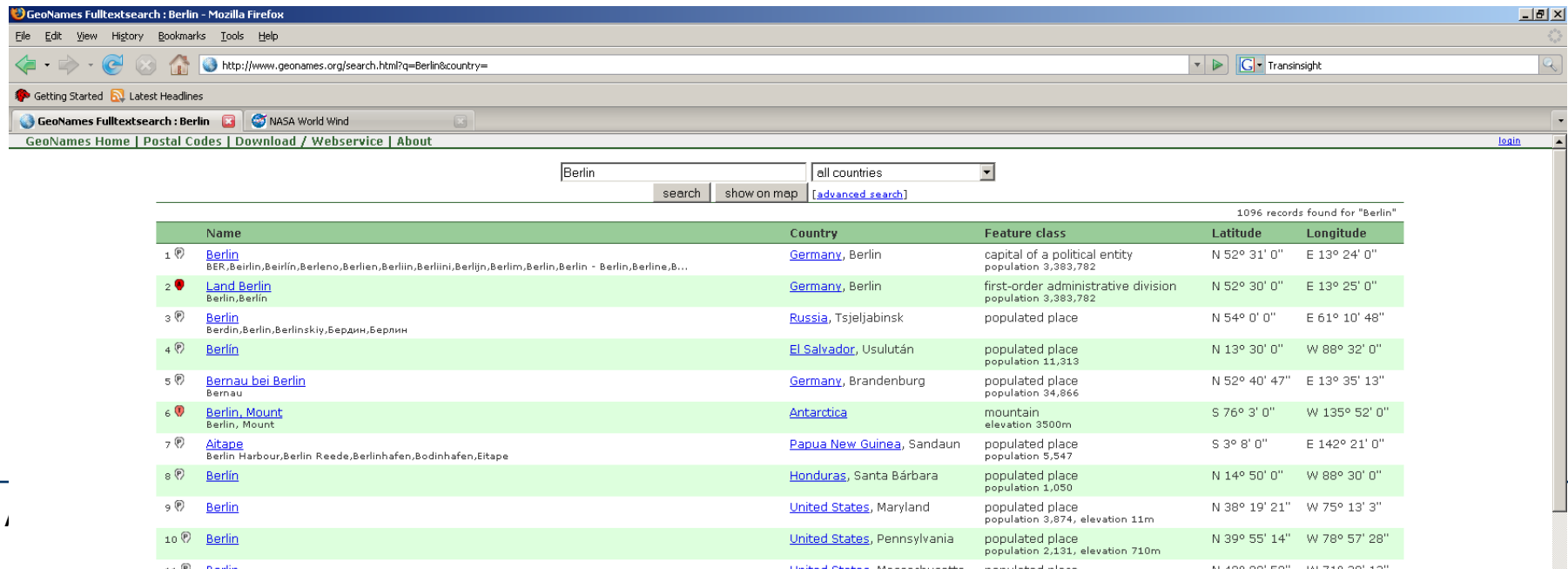
- Nutzung der SIMILE Solvent Screen Scraper
 - Writing screen scrapers can be hard and tedious, that's why you need a tool to help you. Solvent lets you:
 - Interactively highlight parts of the page you wish to scrape, directly in your browser, and obtain the right XPath's for them
 - Inspect the DOM of the captured elements and assign variable name there
 - Automatically generate the javascript code that does the most common features, such as xpath results iterations
 - Choose from different screen scraping templates based on the type of page you are scraping (individual page, multi page, etc..)
 - Edit and execute the scraper code directly in the browser, making the development cycle fast and incremental
 - See the scraped results right in Piggy Bank even without installing the scraper first
 - Save and publish the scraper with the required metadata, so that others can discover it
 - Provide you with all the cheatsheets that you need for javascript, xpath, DOM, RDF and places where you can find RDF vocabularies
- ACM Portal Scraper: http://simile.mit.edu/wiki/ACM_Portal_Scraper
- Upcoming Events Scaper: http://simile.mit.edu/wiki/Ontoworld_Upcoming_Events_Scraper

Mögliche Lösungsskizze (2)

- Benutze SIMILE Piggy Bank um wichtige Geoinformationen wie Veranstaltungsort, sowie weitere Informationen über den Event oder die Publikation für die Abfrage aus dem Semantischen Geoinformationssystem (HTTP Get / Post Anfrage) bereitzustellen
- Features von Piggy Bank
 - Collect information from the Web
 - Save information for future use
 - Tag information with keywords
 - Search & Browse collected information
 - Retrieve saved information
 - Share information you have collected
 - Install screen scrapers
- Hinweis: SIMILE Babel erlaubt Transformationen

Mögliche Lösungsskizze (3)

- Rufe mit den Geoinformationen (Ort) den GeoNames Web Service auf:
 - <http://www.geonames.org/>
 - Liefert weitere Geoinformationen wie Längen- und Breitengrad
 - Hinweis: Aufruf z.B. als im Geoinformationssystem integrierter Web Service Client



The screenshot shows a web browser window displaying the GeoNames search results for 'Berlin'. The search bar contains 'Berlin' and the dropdown menu is set to 'all countries'. The results table shows 1096 records found for 'Berlin'. The table has columns for Name, Country, Feature class, Latitude, and Longitude. The first few results are:

Name	Country	Feature class	Latitude	Longitude
1 Berlin BER, Beirlin, Beirlin, Berleno, Berlien, Berliin, Berliini, Berlijn, Berlim, Berlin, Berlin - Berlin, Berline, B...	Germany , Berlin	capital of a political entity population 3,383,782	N 52° 31' 0"	E 13° 24' 0"
2 Land Berlin Berlin, Berlin	Germany , Berlin	first-order administrative division population 3,383,782	N 52° 30' 0"	E 13° 25' 0"
3 Berlin Berdin, Berlin, Berlinskiy, Бердин, Берлин	Russia , Tsjeljabinsk	populated place	N 54° 0' 0"	E 61° 10' 48"
4 Berlin	El Salvador , Usulután	populated place population 11,313	N 13° 30' 0"	W 88° 32' 0"
5 Bernau bei Berlin Bernau	Germany , Brandenburg	populated place population 34,866	N 52° 40' 47"	E 13° 35' 13"
6 Berlin, Mount Berlin, Mount	Antarctica	mountain elevation 3500m	S 76° 3' 0"	W 135° 52' 0"
7 Aitape Berlin Harbour, Berlin Reede, Berlinhafen, Bodinhafen, Eitape	Papua New Guinea , Sandaun	populated place population 3,547	S 3° 8' 0"	E 142° 21' 0"
8 Berlin	Honduras , Santa Bárbara	populated place population 1,050	N 14° 50' 0"	W 88° 30' 0"
9 Berlin	United States , Maryland	populated place population 3,874, elevation 11m	N 38° 19' 21"	W 75° 13' 3"
10 Berlin	United States , Pennsylvania	populated place population 2,131, elevation 710m	N 39° 55' 14"	W 78° 57' 28"

Mögliche Lösungsskizze (4)

- Darstellung der Event und Publikationsinformationen mit Hilfe der Geokoordinaten auf der 3D Weltkugel von NASA World Wind
 - <http://worldwind.arc.nasa.gov/>
- als Java Anwendung (schöner Java Web Start, Applet oder Integration als View in SIMILE Exhibit)



Hinweis zur Gruppenaufgabe

- 4er Teams
- Aufteilung der Aufgabe im Team beliebig, z.B.
 - Anbindung an Piggy Bank ACM und Event Screen Scraper / Bereitstellung der Daten für den GeoNames Client
 - Implementierung des Clients zum Aufruf des GeoNames Services / Bereitstellung der Ortskoordinaten und Schnittstellen für die Integration in das Geoinformationssystem
 - Implementierung des grafischen Geoinformationssystem
 - Implementierung des Anwendungslogik für die Darstellung der Event und Publikationsinformationen auf dem NASA World Wind 3D Globus (+ Konfigurationsmöglichkeiten für Nutzer / Nutzerschnittstelle)
 - Integration der SIMILE Screen Scraper und GeoNames Client (und Implementierung des Java Webstarts oder als Web Applikation unter Tomcat oder Erweiterung von SIMILE Exhibit)
- Wichtig:
 - Schnittstellen (Interfaces) zwischen den Arbeitspaketen im Team frühzeitig festlegen
- Eigene Ideen und Lösungen
 - Toolunterstützung ausnützen

Übung 10: Was noch?

- **Gruppenaufgabe**
 - 4er Teams
- **Termine**
 - Ausgabe: 3.2.2009 Abgabe bis: 10.2.2009, 16:00 Uhr
- Fragen?

Viel Spaß und Erfolg !

