

LinkedGeoData

Eine räumliche Dimension für das Datenweb

Sören Auer, Jens Lehmann, Claus Stadler

AKSW, Universität Leipzig



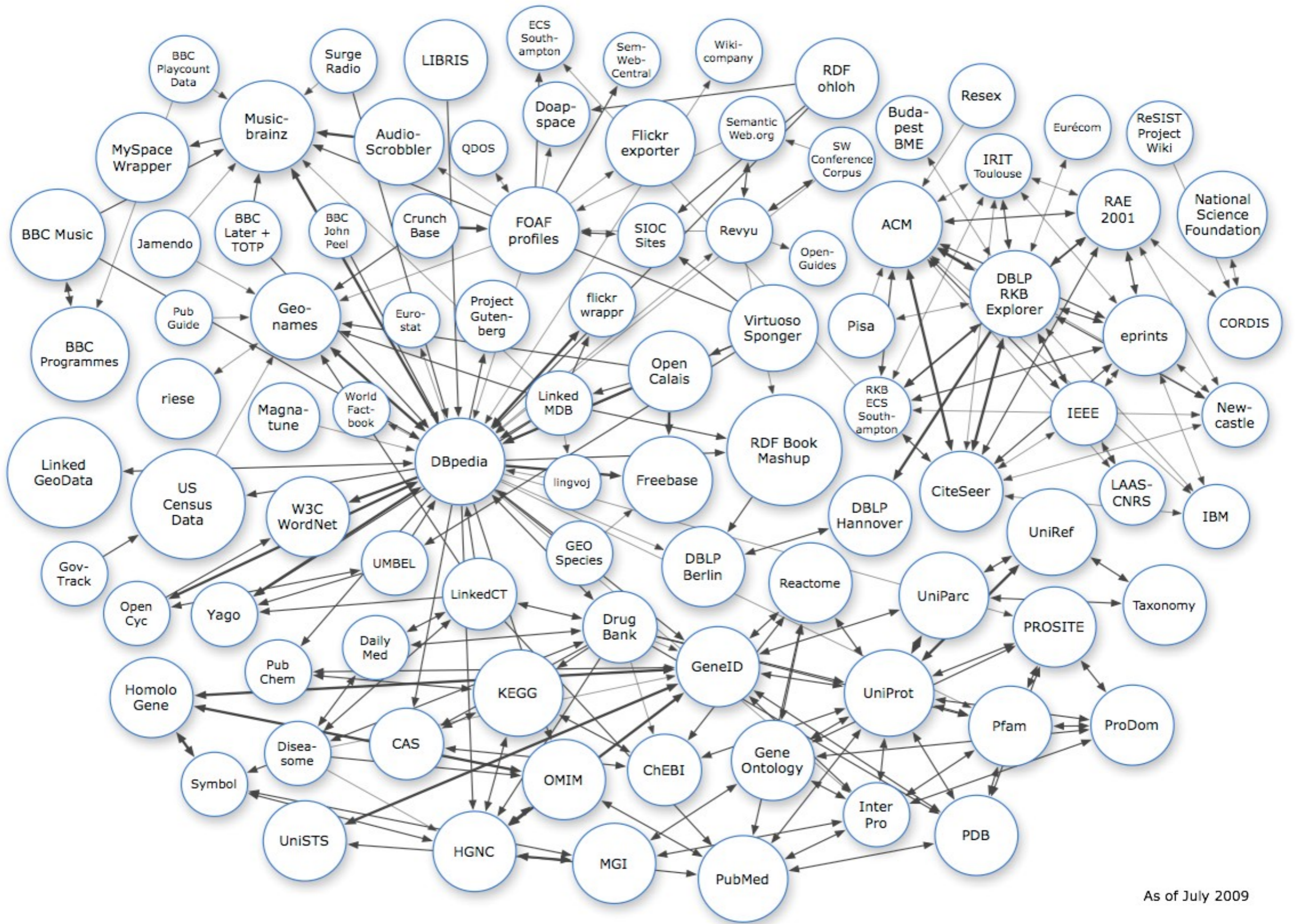
Worum geht's?

Aufbereitung und Veröffentlichung von **OpenStreetMap**
Daten nach **Semantic Web** Prinzipien

(RDF, LinkedData, SPARQL Endpunkt)

Motivation

- Was sind die Angebote von Bäckereien in der Nähe?
- Wo innerhalb Leipzig hat ALDI seine Filialen?
- Was sind die historischen Sehenswürdigkeiten entlang einer Radstrecke?
- Wo ist die nächste Müllsammelstelle?



<http://richard.cyganiak.de/2007/10/lod/>

Linking Open Data - Dataset Cloud

- Enthält Datensets mit **räumlichen Features**

GeoNames, DBpedia, US consenus, EuroStat, ...

- Beschränkt auf **populäre/große Entitäten**

Länder, berühmte Plätze, etc

- Es **fehlen**

Gebäude, Straßen, Briefkästen, Müllsammelstellen, ...

Daher

Wenn es für diese **Entitäten** (stabile) **URLs** gäbe...

- ...Könnten sowohl

beim **Abruf** der URLs **Links** auf andere **Wissenbasen** zurückgeliefert werden,

- als auch

Andere Wissenbasen diese Entitäten **referenzieren**

Daher

Mächtigeren Applikationen durch mächtigere Datensätze ermöglichen

LinkedGeoData soll im Semantic Web ähnliche Rolle wie DBpedia spielen

OpenStreetMap (OSM)

<http://www.openstreetmap.org/>

- Grosses Community-Projekt mit dem Ziel der Erstellung einer **freien Karte der Welt**
- Informationsdichte und -qualität **übertrifft stellenweise die von kommerziellen Anbietern**



<http://www.vimeo.com/2598878>

OSM Datenmodell

- **Knoten**: Latitude, Longitude
- **Wege**: Menge von Knoten
- **Relationen**: Gruppierungen von Knoten und Wegen

- **+ Tags**: Schlüssel/Wert Paare

Datesatzbeispiel

```
lgd=# select node_id, k, v from node_tags limit 10;
```

```
node_id | k | v
```

```
-----+-----+-----
```

```
687139839 | amenity | waste_basket
```

```
687139860 | crossing | uncontrolled
```

```
687139860 | highway | crossing
```

```
687139871 | barrier | entrance
```

```
687139871 | foot | yes
```

```
687139884 | name | Yellow
```

```
687139884 | shop | shoes
```

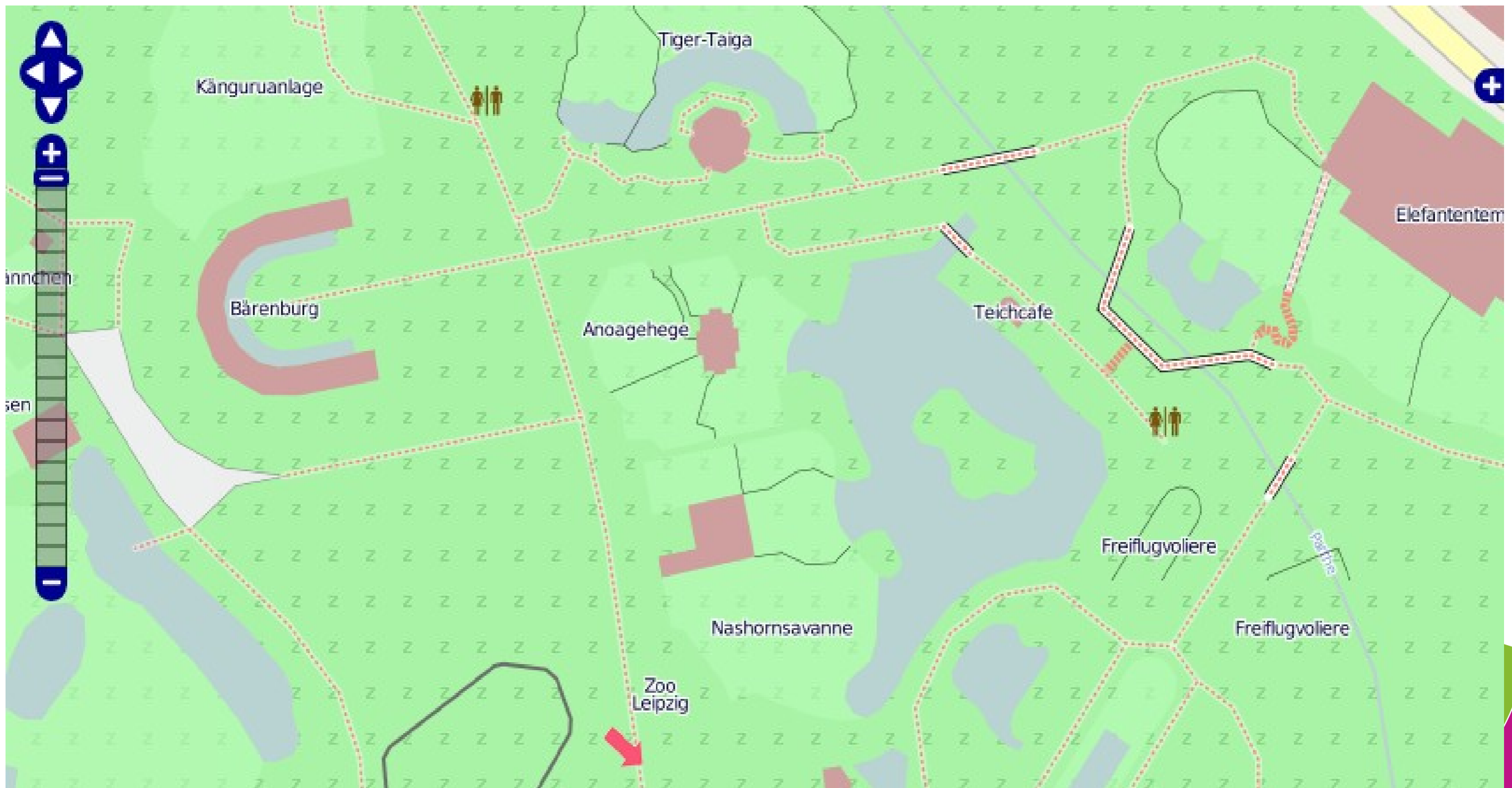
```
687139885 | amenity | waste_basket
```

```
687139889 | amenity | waste_basket
```

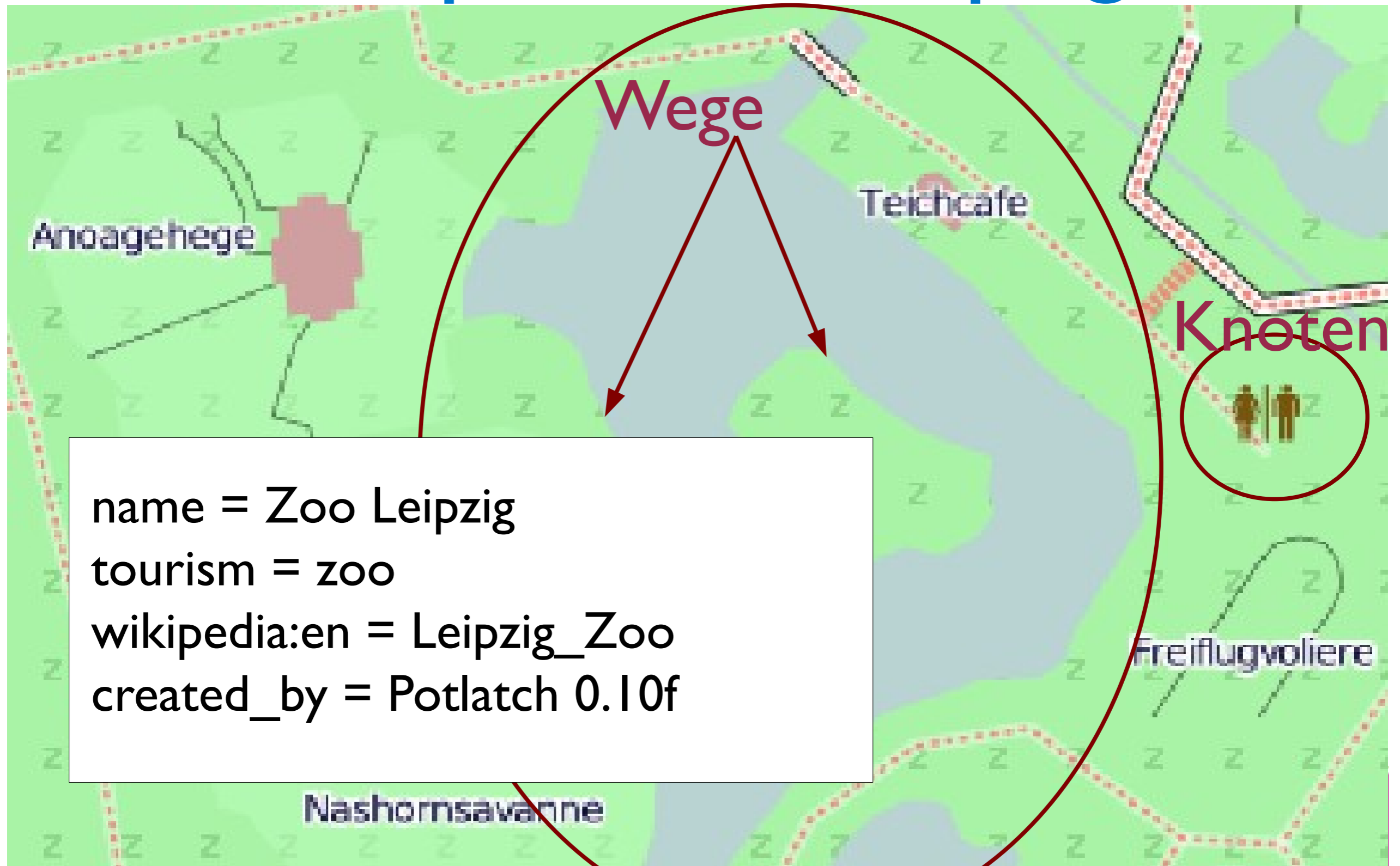
```
687139908 | shop | supermarket
```

(10 Zeilen)

Beispiel: Zoo Leipzig



Beispiel: Zoo Leipzig



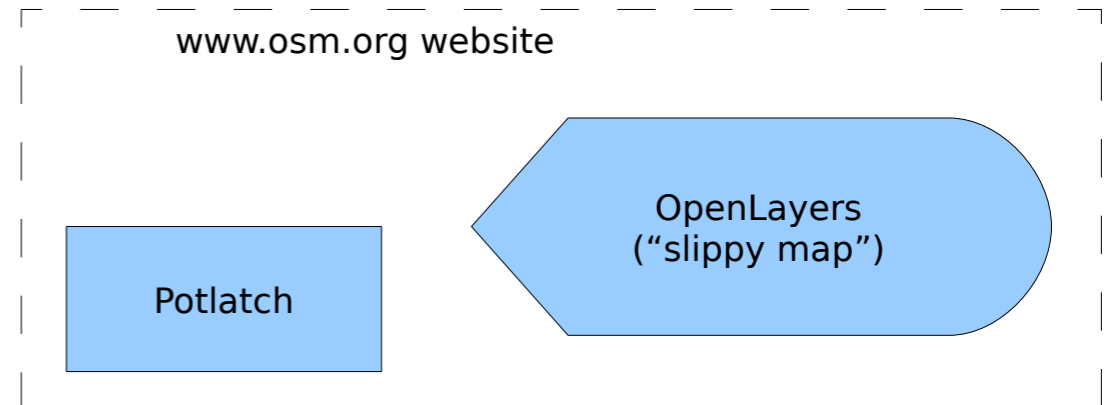
name = Zoo Leipzig
tourism = zoo
wikipedia:en = Leipzig_Zoo
created_by = Potlatch 0.10f

OSM Statistik

Kategorie	Gesamt Juni 2009	Gesamt April 2010
users	127 tausend	261 tausend
Uploaded GPS points	915 millionen	1.5mrd
nodes	374 millionen	600 millionen
ways	30 millionen	48 millionen

(Credits to Iván Sánchez Ortega for this overview.)

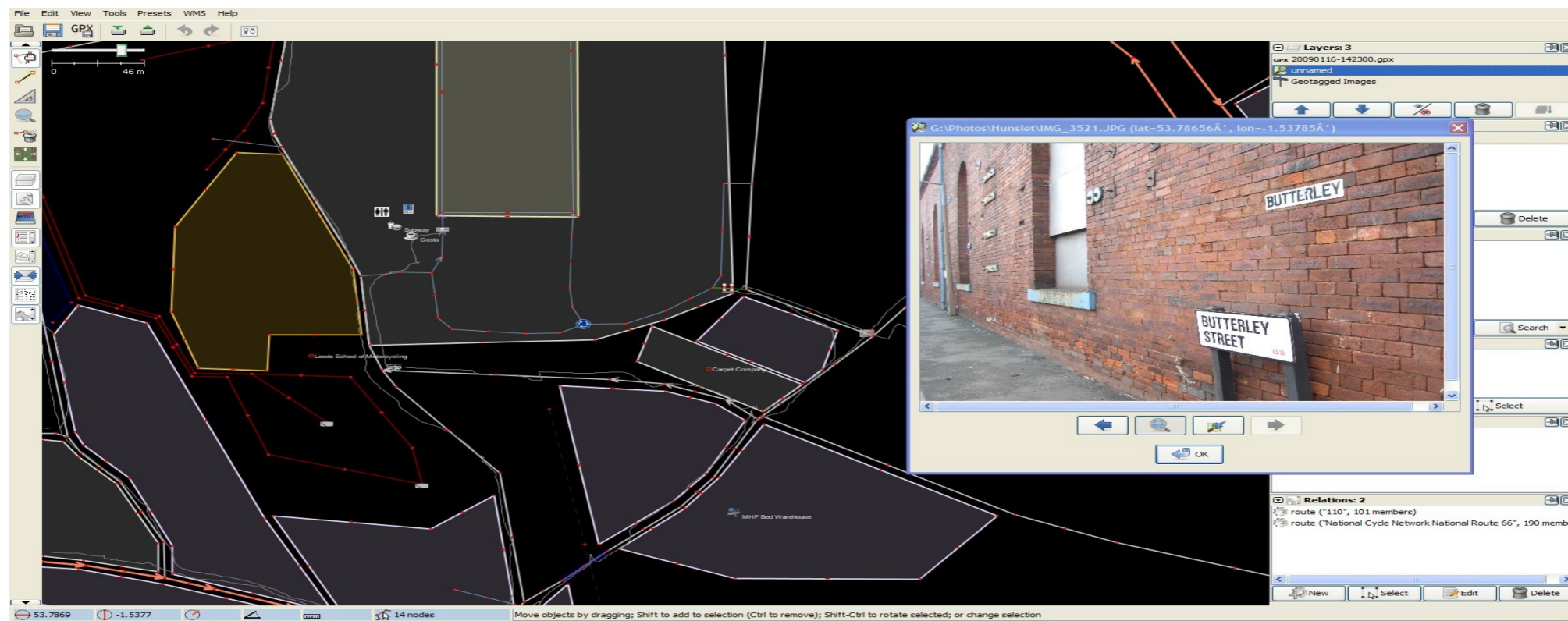
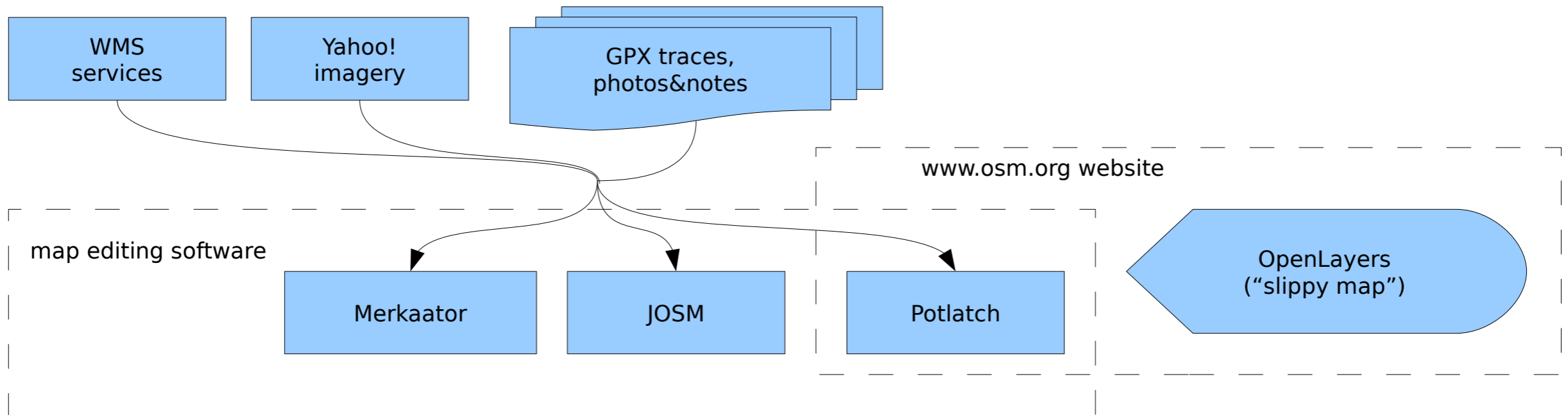
OSM: Website



The screenshot shows the OpenStreetMap website interface. At the top, there are navigation tabs: "Karte", "Bearbeiten", "Chronik", "Export", "GPS-Tracks", and "Blogs". On the right side, there are links for "Anmelden" and "Registrieren". The main content is a map of the Washington, DC area, showing roads, airports, and various islands. On the left side, there is a sidebar with several sections: "OpenStreetMap" with a magnifying glass icon and the text "Die freie Wiki-Weltkarte"; a paragraph describing OpenStreetMap as a free, editable world map; a section for "Hilfe & Wiki", "News-Blog", "Shop", and "Legende"; a search bar with the text "Suchen" and "Wo bin ich?"; a "Spenden" button; and a Creative Commons license icon (CC BY SA). At the bottom right of the map, there are links for "Permalink" and "Shortlink".

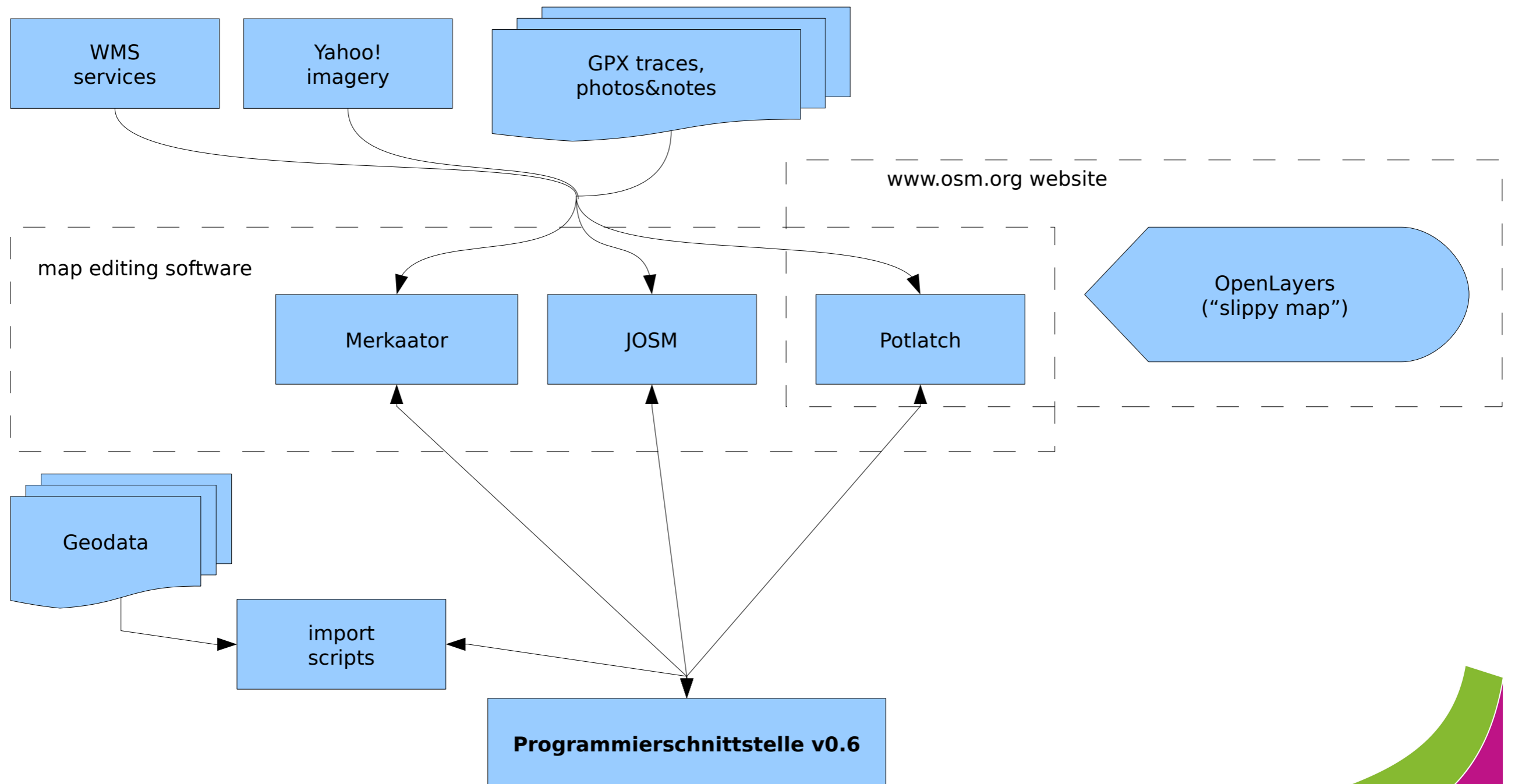
(Credits to Iván Sánchez Ortega for this overview.)

OSM: Editoren



(Credits to Iván Sánchez Ortega for this overview.)

OSM: Schnittstelle



LinkedGeoData

- Konvertierung der OSM Daten nach **RDF/OWL**
- Neue **Links** zu **DBpedia** berechnet
- Browser

RDF = Resource Description Framework

OWL = Web Ontology Language

Konvertierung nach RDF/OWL

- Knoten, Wege und Kanten

Subklassen von `wgs84:SpatialThing`

- Positionsangaben

`wgs84:{long, lat}`, `georss:{point, line, polygon}`

Konvertierung nach RDF/OWL

- Tags fallen in 3 **Kategorien**
 - Klassifikationsattribute, z.B. highway = motorway
 - Beschreibungsattribute, z.B. created_by = JOSM
 - Datenattribute, z.B. Postcode = 01187
- **Konvertiert** nach
 - 500 Klassen
 - 50 Object-Properties
 - 15000 Datentyp-Properties

Konvertieren von Tags

- Drei allgemeine Mappings

- Text

(5, name, München) → lgdn:5#id lgdv:name "München"

(5, name:en, Munich) → lgdn:5#id lgdv:name "Munich"@en

- Datentypen

(6, seats, 5) → lgdn:6#id lgdv:seats "5"^^xsd:integer

- Klassen

(7, amenity, embassy) → lgdn:7 a lgdv:embassy

Konvertieren von Tags

- Spezial Mappings

- (wikipedia:en, Leipzig) →

lgdv:wikiLink <http://en.wikipedia.org/wiki/Leipzig>

(anstatt lgdv:wikipedia "Leipzig"@en)

- Viele weitere Tags definiert in der **OSM Ontologie**

http://wiki.openstreetmap.org/wiki/Map_features

DBpedia-LGD Mapping

- Idee:

Durch **Verlinken** mit DBpedia wird LGD indirekt auch mit **anderen Wissensbasen** vernetzt.

DBpedia-LGD Mapping

Etliche Entitäten in DBpedia haben GeoCoordinaten

`rdfs:label`

■ London

`geo:lat`

■ 51.508057 (xsd:float)

`geo:long`

■ -0.124722 (xsd:float)

<http://dbpedia-live.openlinksw.com/page/London>

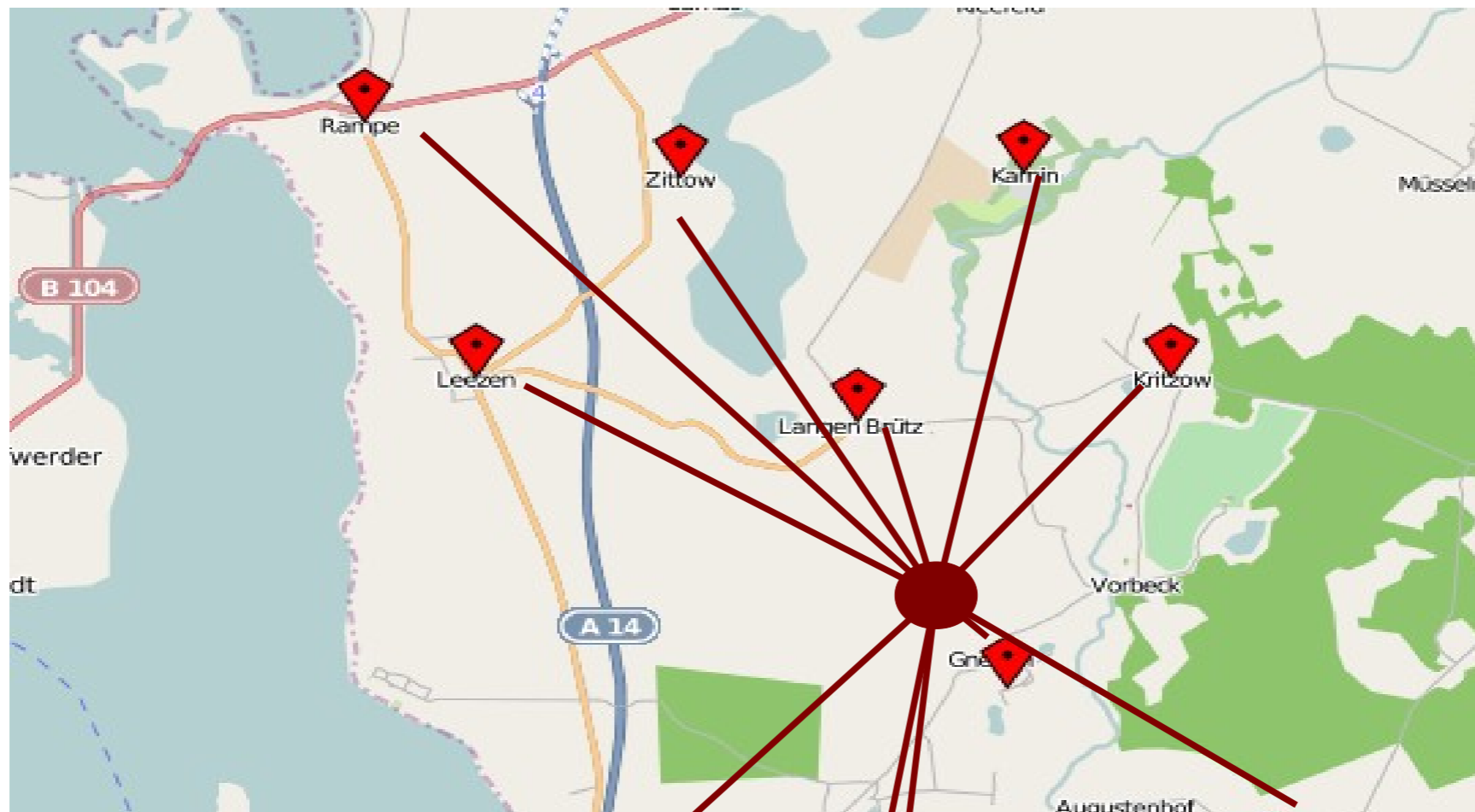


DBpedia-LGD Mapping

Für jeden **Punkt** mit Geo-Koordinaten aus **DBpedia**, **suche** passende Kandidaten in **LGD** basierend auf:

- Typ Information (Schema Matching)
- Räumliche Distanz
- Namensähnlichkeit

LGD-DBpedia Mapping



Ab einem **Grenzwert** wird der **höchstbewertetste** LGD-Kandidat-Punkt mit **owl:sameAs** zur DBpedia-Resource verlinkt



Ergebnis

- 50.000+ Übereinstimmungen
- Decken 53.8% der DBpedia Geo-Entitäten ab

Type	#Matches	Rate
city	45729	70.9%
railway station	929	24.8%
university	210	13.3%
school	1483	38.4%
airport	649	8.4%
lake	1014	22.1%
country	160	20.1%
island	313	29.8%
mountain	1475	24.5%
river	677	32.0%
lighthouse	25	4.3%
stadium	346	17.0%

- Gründe fürs Fehlschlagen:

- Resource existiert nicht in LGD
- Fehlerhafte/Unzureichende

Beschreibung in DBpedia/LGD

Zugriff auf die Daten

REST-Interface

latitude longitude radius

<http://LinkedGeoData.org/triplify/near/48.213,16.359/1000>

Zugriff auf die Daten

<http://linkedgedata.org/triplify/near/48.213,16.359/1000/class/pub?format=n3>

```
<http://linkedgedata/triplify/node/334390933#id>
  a      <http://linkedgedata.org/vocabulary#pub> ;
  <http://linkedgedata.org/vocabulary#addr:city>
    "Wien" ;
  <http://linkedgedata.org/vocabulary#addr:country>
    "AT" ;
  <http://linkedgedata.org/vocabulary#addr:housenumber>
    "7" ;
  <http://linkedgedata.org/vocabulary#addr:postcode>
    "1010" ;
  <http://linkedgedata.org/vocabulary#addr:street>
    "Naglergasse" ;
  <http://linkedgedata.org/vocabulary#created_by>
    "xybot" ;
  <http://linkedgedata.org/vocabulary#name>
    "Bockshorn Irish Pub" ;
  <http://linkedgedata.org/vocabulary#note>
    "Wiens ältestes Pub" ;
  <http://linkedgedata.org/vocabulary#website>
    "http://www.bockshorn.at" .
```

Browser Demo

Instanzen

Zoomstufen

RDF Export

and was developed by [AKSW research group](#).

LinkedGeoData.org

Facets

Instances

Search: powered by [Namefinder](#)

Link | RDF-Export

Class hierarchy

- [+] [amenity](#)(19)
- [+] [highway](#)(12)
- [+] [leisure](#)(1)
- [+] [man_made](#)(1)
- [+] [railway](#)(1)
- [+] [shop](#)(4)

Properties

- [+] [cuisine](#)(1)
- [+] [denomination](#)(2)
- [+] [religion](#)(2)

Query took 14ms.

1. amenity: recycling

2. **Club 11**
amenity: pub

3. man_made: MDF

4. amenity: bicycle_parking

5. highway: traffic_signals

6. highway: traffic_signals

7. highway: traffic_signals

8. **Reichenbachstraße**
highway: bus_stop
railway: tram_stop

9. **Hl. Simeon vom Wunderbaren Berge (Russ.-orthodoxe Kirche)**
amenity: place_of_worship
denomination: orthodox
religion: christian

10. highway: traffic_signals

11. amenity: recycling

12. **Reichenbachstraße**
highway: bus_stop

13. amenity: post_box

14. amenity: telephone

15. highway: traffic_signals

16. amenity: recycling

17. highway: traffic_signals

18. highway: turning_circle

19. highway: traffic_signals

20. **Lukaskirche**
amenity: place_of_worship
denomination: lutheran
religion: christian

21. amenity: parking

22. **BÄckerei Möbius**
shop: bakery

23. shop: butcher

24. shop: bakery

25. **Netto**

100 m
200 ft

13.72846, 51.03480

Facetten

Zukünftige Arbeit

- **Datenqualität** verbessern / Fehler ausbessern

- Integration der **OSM-Ontologie**

http://wiki.openstreetmap.org/wiki/Map_features

- **Verlinkung** mit anderen Wissensbasen

GeoNames, FishBase, SeaLifeBase, AquaMaps

- Mehr Labels + **Internationalisierung**

<http://translatewiki.net/wiki/Translating:OpenStreetMap>

- Wege und Relationen berücksichtigen

Momentan noch Fokus auf Knoten

Zukünftige Arbeit

- **LinkedData**-Schnittstelle
 - **Performance** verbessern
 - Mehr **Funktionalität**
- **SPARQL** Endpunkt anbieten

OpenLink stellt dafür kommerzielle **Virtuoso** Version zur Verfügung

- **Live-Synchronisation**

Zukünftige Arbeit

- Browser verbessern
 - Editier-Modus + Rücktransformation von Tripel auf Tags
 - ...

Zukünftige Arbeit

Daten und Erkenntnisse an die OSM Community zurückfliessen lassen.

Danke für Eure Aufmerksamkeit!



Kaffee!

