

# OpenStreetMap und Open Government Data in PostGIS

Offene geografische Daten zusammenführen, verarbeiten und  
visualisieren

Balázs Bárány

Rapid-I GmbH

Linuxwochen Wien 2013

# Inhalt

Überblick und Grundlagen

PostgreSQL und PostGIS

Einlesen der OSM-Daten

Arbeit mit Geodaten in PostGIS

- Einfache Abfragen

- Abfragen mit geografischen Funktionen

- Geodaten-Feld im Ergebnis

- Views und Tabellen

Open Data importieren

- Open Government Data in Österreich

- Datenimport mit RapidMiner

# Geodaten in der Datenbank

- ▶ Warum Geodaten in der Datenbank speichern?

# Geodaten in der Datenbank

- ▶ Warum Geodaten in der Datenbank speichern?
- ▶ Warum PostGIS?

# Wichtige Begriffe

- ▶ Arten von Geodaten
  - ▶ Rasterdaten
  - ▶ Vektordaten

# Wichtige Begriffe

- ▶ Arten von Geodaten
  - ▶ Rasterdaten
  - ▶ Vektordaten
- ▶ Abbildung der Erde in zwei Dimensionen
  - ▶ Projektion
  - ▶ Koordinatenbezugssystem

# Österreich in falscher und richtiger Projektion



Abbildung: WGS 84

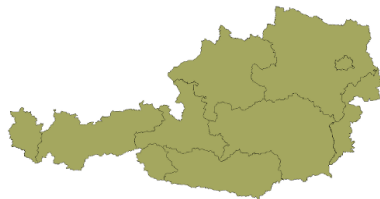


Abbildung: ETRS89 / Austria Lambert

# Geografische Objekte

- ▶ Punkt
- ▶ Linie



# Geografische Objekte

- ▶ Punkt
- ▶ Linie
- ▶ Fläche (Polygon)
  - ▶ Multipolygon

# Dateiformate für Geodaten

- ▶ ESRI Shapefile

# Dateiformate für Geodaten

- ▶ ESRI Shapefile
- ▶ OpenStreetMap

# Dateiformate für Geodaten

- ▶ ESRI Shapefile
- ▶ OpenStreetMap
- ▶ GPX, KML

# Dateiformate für Geodaten

- ▶ ESRI Shapefile
- ▶ OpenStreetMap
- ▶ GPX, KML
- ▶ Comma separated values

# Dateiformate für Geodaten

- ▶ ESRI Shapefile
- ▶ OpenStreetMap
- ▶ GPX, KML
- ▶ Comma separated values
- ▶ Konvertierungswerkzeuge
  - ▶ OGR
  - ▶ GPSTabel

# OpenStreetMap

- ▶ Community

# OpenStreetMap

- ▶ Community
- ▶ Inhalte



# OpenStreetMap

- ▶ Community
- ▶ Inhalte
- ▶ Lizenz: Open Data Commons Open Database Licence

# PostgreSQL-Installation und Konfiguration

- ▶ Linux: Distributionspakete oder Binärpakete vom PostgreSQL-Projekt
- ▶ Windows: Installer

# PostgreSQL-Installation und Konfiguration

- ▶ Linux: Distributionspakete oder Binärpakete vom PostgreSQL-Projekt
- ▶ Windows: Installer
- ▶ Netzwerkzugriff freischalten
- ▶ Zugriffsrechte: User und Gruppen

# PostGIS-Einrichtung: Template-Datenbank

## Beispiel

```
CREATE DATABASE gis_template;  
\connect gis_template;  
\i /pfad/zu/postgis.sql  
\i /pfad/zu/spatial_ref_sys.sql  
\i /pfad/zu/../../postgis_comments.sql  
CREATE EXTENSION hstore;
```

# PostGIS: neue geografische Datenbank

## Beispiel

```
CREATE DATABASE gis WITH TEMPLATE gis_template  
OWNER gis;
```

# PostGIS: neue geografische Datenbank

## Beispiel

```
CREATE DATABASE gis WITH TEMPLATE gis_template  
OWNER gis;
```

- ▶ geometry\_columns
- ▶ spatial\_ref\_sys

# osm2pgsql

- ▶ 64-bit-Betriebssystem,  $\geq 8$  GB!

# osm2pgsql

- ▶ 64-bit-Betriebssystem,  $\geq 8$  GB!
- ▶ Download von Geofabrik



# osm2pgsql

- ▶ 64-bit-Betriebssystem,  $\geq 8$  GB!
- ▶ Download von Geofabrik
- ▶ Skript auf der Vortragshomepage

# OSM-Daten in PostGIS

- ▶ Basistabellen: `osm_austria_lines`, `osm_austria_point`,  
`osm_austria_polygon`
- ▶ `osm_austria_roads`

# OSM-Daten in PostGIS

- ▶ Basistabellen: `osm_austria_lines`, `osm_austria_point`, `osm_austria_polygon`
- ▶ `osm_austria_roads`
- ▶ Einträge in `geometry_columns`

# Bundesländer

## Beispiel

```
SELECT DISTINCT name
FROM osm_austria_polygon
WHERE boundary='administrative'
      AND admin_level = '4'
ORDER BY name;
```

name
Bayern
Burgenland
Kärnten
Niederösterreich
Oberösterreich
Salzburg
Steiermark
Tirol
Vorarlberg
Wien

# Häufigste Namen von Restaurants

## Beispiel

```
SELECT name, COUNT(osm_id)
AS anzahl
FROM osm_austria_point
WHERE amenity = 'restaurant'
AND name <> ''
GROUP BY name
ORDER BY anzahl DESC
LIMIT 10
```

name	anzahl
Kirchenwirt	43
Gasthof zur Post	16
Pizzeria	13
Gasthaus	11
Dorfwirt	10
Interspar Restaurant	9
Asia	9
Pizzeria Venezia	9
Mensa	8
Gasthof Post	7

# Öffi-Haltestellen in der Nähe

## Beispiel

```
WITH fh_technikum AS (  
  SELECT way FROM osm_austria_polygon  
  WHERE name = 'FH Technikum Wien'  
)  
SELECT DISTINCT p.name,  
  CASE WHEN p.highway = 'bus_stop' THEN 'Bushaltestelle'  
        WHEN p.railway = 'subway_entrance' THEN 'U-Bahn-Station'  
        ELSE 'Straßenbahnhaltestelle'  
  END AS art,  
  ROUND(ST_Distance(p.way, fh.way)) AS "Entfernung (m)"  
FROM osm_austria_point p  
INNER JOIN fh_technikum fh ON ST_Distance(p.way, fh.way) < 250  
WHERE (highway = 'bus_stop' OR railway IN ('subway_entrance', 'tram_stop'))  
ORDER BY ROUND(ST_Distance(p.way, fh.way));
```

# Ergebnis: Öffi- Haltestellen

name	art	Entfernung (m)
Höchstädtplatz	Straßenbahnhaltestelle	34
Höchstädtplatz	Straßenbahnhaltestelle	64
Höchstädtplatz	Bushaltestelle	74
Höchstädtplatz	Bushaltestelle	89
Höchstädtplatz	Straßenbahnhaltestelle	113
Dresdner Straße	Straßenbahnhaltestelle	173
Dresdner Straße	Bushaltestelle	181
Dresdner Straße	Bushaltestelle	199
Dresdner Straße (Meldemannstraße)	Bushaltestelle	228
Pasettistraße	U-Bahn-Station	235
Dresdner Straße	U-Bahn-Station	237

# Koordinaten des Riesenrades

## Beispiel

```
SELECT name, way, ST_AsText(way) AS fläche  
FROM osm_austria_polygon  
WHERE name = 'Riesenrad';
```

name	way	fläche
Riesenrad	01030000...08C1D41	POLYGON(((627460.39 484102,...))

(way und fläche abgekürzt)



# View für die Bundesländer

## Beispiel

```
CREATE VIEW osm_bundesländer_namen AS
SELECT DISTINCT name
FROM osm_austria_polygon
WHERE admin_level = '4'
      AND boundary='administrative'
      AND name NOT IN ('Bayern')
ORDER BY name;
```

# Bundesländer-Tabelle mit Geodaten

## Beispiel

```
CREATE TABLE osm_bundesland AS
SELECT MIN(osm_id) AS id, name, ST_Collect(way) AS geo
FROM osm_austria_polygon
WHERE admin_level = '4'
      AND boundary='administrative'
      AND name NOT IN ('Bayern')
GROUP BY name
ORDER BY name;
```

## Neue Tabelle/View registrieren

- ▶ Erfassung der neuen Tabelle (View) in `geometry_columns`

### Beispiel

```
SELECT  
Populate_Geometry_Columns('osm_bundesland'::regclass);
```

## Neue Tabelle/View registrieren

- ▶ Erfassung der neuen Tabelle (View) in `geometry_columns`

### Beispiel

```
SELECT  
Populate_Geometry_Columns('osm_bundesland'::regclass);
```

- ▶ Notwendig für die meisten Clients
- ▶ Überprüfungen, Konsistenzregeln automatisch hinzugefügt

# Beispiele geografischer Funktionen

- ▶ ST\_Area
- ▶ ST\_Transform
- ▶ ST\_Buffer
- ▶ ST\_GeomFromText

# Open Government Data in Österreich

- ▶ data.gv.at
  - ▶ Portal zu den Open-Data-Seiten öffentlicher Institutionen
  - ▶ Suchfunktion

# Open Government Data in Österreich

- ▶ data.gv.at
  - ▶ Portal zu den Open-Data-Seiten öffentlicher Institutionen
  - ▶ Suchfunktion
- ▶ (statistik.at)

# RapidMiner

- ▶ Open-Source-Software für Data Mining und analytische Datenintegration



# RapidMiner

- ▶ Open-Source-Software für Data Mining und analytische Datenintegration
  - ▶ Datenanalyse, Visualisierung, Vorhersagen

# RapidMiner

- ▶ Open-Source-Software für Data Mining und analytische Datenintegration
  - ▶ Datenanalyse, Visualisierung, Vorhersagen
  - ▶ Datenimport, Verarbeitung, Export

# RapidMiner

- ▶ Open-Source-Software für Data Mining und analytische Datenintegration
  - ▶ Datenanalyse, Visualisierung, Vorhersagen
  - ▶ Datenimport, Verarbeitung, Export
- ▶ Leicht zu bedienende GUI

# RapidMiner

- ▶ Open-Source-Software für Data Mining und analytische Datenintegration
  - ▶ Datenanalyse, Visualisierung, Vorhersagen
  - ▶ Datenimport, Verarbeitung, Export
- ▶ Leicht zu bedienende GUI
- ▶ Automatisierung
  - ▶ Befehlszeile
  - ▶ RapidAnalytics

# RapidMiner-Oberfläche

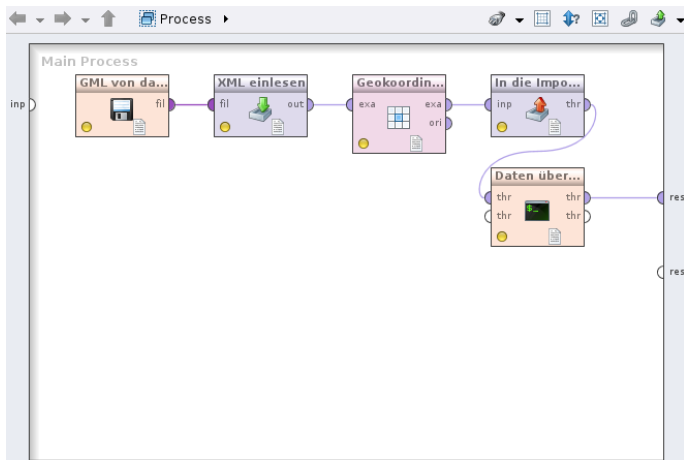
The screenshot displays the RapidMiner 5.3.008 interface. The main workspace shows a workflow titled "Main Process" with the following steps:

- GML von da...**: Reads GML data from a source.
- XML einlesen**: Parses the XML data.
- Geokoordin...**: Processes the geographic coordinates.
- In die Impo...**: Imports the data into a database.
- Daten über...**: A subsequent process, possibly for data verification or reporting.

The interface includes several panels:

- Overview**: Shows a high-level view of the process flow.
- Repositories**: Lists data sources like "Samples (none)", "DB", "repo 5.0 (lib)", and "Berta (Entwicklungs-RA) (offline)".
- Operators**: A library of various data processing operators such as "Process Control", "Parameter", "Loop", "Branch", "Collections", "Multiply", "Handle Exception", "Remember", "Recall", "Join Paths", "Throw Exception", "Utility", "Macros", and "Logama".
- Parameters**: Configurable settings for the current process, including "logverbosity", "logfile", "resultfile", "random seed", "send mail", and "encoding".
- Problems**: A log window showing "No problems found".
- Help**: A section with a "Synopsis" and "System Monitor" showing system resources like "Mem: 1,3 GB" and "Total: 26 MB".

# Datenimport mit RapidMiner



# Ausblick

- ▶ Publikation freier Karten

# Ausblick

- ▶ Publikation freier Karten
- ▶ Interaktive Karten im Web



# Ausblick

- ▶ Publikation freier Karten
- ▶ Interaktive Karten im Web
- ▶ Analytik

# Fragen?

Anleitung: <http://tud.at/linuxwochen/2013-osm-postgis/>

Scan:

