

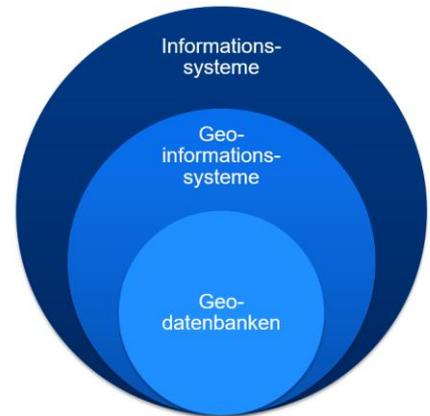
OS Aktuelle Datenbanktrends

Geodatenbanken

Geo-Informationssysteme

Definition

Geo-Informationssysteme (GIS) sind Informationssysteme, mit denen „[...] raumbezogene Daten digital erfasst und redigiert, gespeichert und reorganisiert, modelliert und analysiert sowie alphanumerisch und graphisch präsentiert werden,“ (Bill, 1994)



Anwendungsbereiche

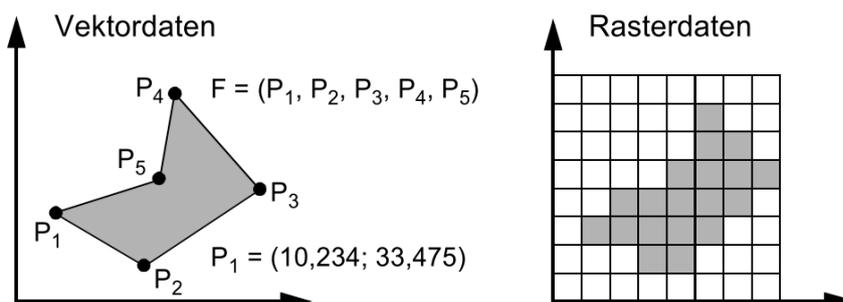
- Verkehr: Navigationssysteme, Verkehrsleitsysteme
- Logistik: Routen von Paketzustellern
- Katasterwesen: (automatisches) Führen von Liegenschaftskarten
- Geomarketing: standortbasierte Angebote
- Umwelt-Monitoring: Beobachtung von bedrohten Tierartsbeständen
- Umweltschutz: Gewässerverschmutzung, Analyse von Umweltschäden
- Amtliche Statistik: Bevölkerungs- / Kriminalstatistik

Geodaten

Eigenschaften

Geometrische Eigenschaften

- Lage und Ausdehnung von Objekten im Raum
- Darstellung als Vektoren oder im Raster

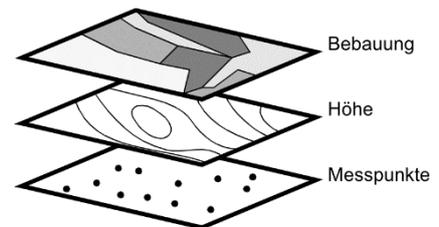


Topologische Eigenschaften

- Relative räumliche Beziehungen von Geoobjekten
- Bsp.: Nachbarschaft, Enthaltensein, Überschneidung
- Können implizit errechnet oder explizit angegeben werden

Thematische Eigenschaften

- entsprechen Sachattributen
- Beispiele:
 - nominal z.B. Bezeichnungen
 - qualitativ, z.B. Wochentag der Müllabfuhr
 - quantitativ, z.B. Einwohnerzahl



Temporalen Eigenschaften

- (Gültigkeits-) Zeitpunkte bzw. Zeiträume für Geoobjekte
- Abbildung von Dynamiken bewegter Geoobjekte
- Bsp.: Wahlergebnisse verschiedener Bundesländer im zeitlichen Verlauf

Standardisierung von Geodaten

Organisationen

- Open Geospatial Consortium (OGC)
- Technisches Komitee 211 „Geographic Information/ Geomatics“ (ISO/TC 211)
- Seit 1998 Zusammenarbeitsabkommen



Wichtige Spezifikationen

Simple-Feature-Modell

- ISO 19125-1 „Simple Feature Access – Common Architecture“ (2004)
- ISO 19125-2 „Simple Feature Access – SQL Option“ (2004)
- beschreibt zweidimensionale Vektor-Geometrien
- definiert (neben Datentypen) räumliche Operationen

Simple-Features

ISO 19125: Geometrien im zweidimensionalen Raum, deren Stützpunkte geradlinig miteinander verbunden sind

SQL/MM Spatial

- ISO/IEC 13249-3 SQL/MM Spatial
- spezifiziert Datenmodell für Geodaten
- Anwendung: SQL Erweiterungen
- große Ähnlichkeit zum Simple-Feature-Modell
- Unterschiede zum Simple Feature Modell:
 - Netzwerk- und Topologiedatenbankschema
 - 3D Funktionen
 - zusätzlich Geometrien für Kreisbögen

Funktionen von Geodatenbanksystemen

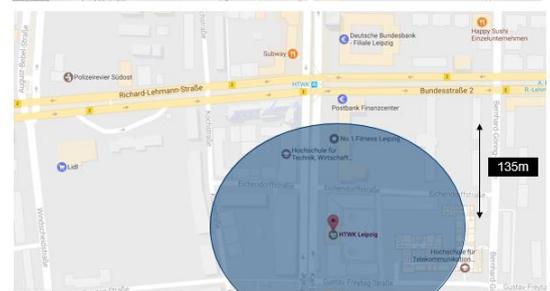
Distance

- Berechnet geringste Entfernung zwischen zwei Geometrien



Buffer

- Ermittelt für vorgegebenen Abstand eine Pufferzone um eine Geometrie



ConvexHull

- Berechnet konvexe Hülle einer Geometrie



Intersection

- Berechnet Schnitt zwischen zwei Geometrien



Difference

- Berechnet die Differenz zweier Geometrien



SymDifference

- Vereinigung beider Differenzen zwischen Geometrie A und B (XOR)



Union

- Ermittelt Vereinigung zweier Geometrien



Topologische Prädikate

- Boolesche Aussagen über topologische Beziehungen zweier Geometrien
- Ermittlung Semantik zweier Prädikate mithilfe des **9-Intersection-Modells (9IM)**
- Aufteilung von Geometrien:
 - Rand
 - Inneres
 - Äußeres

A disjoint B	$\begin{bmatrix} \text{false} & \text{false} & \text{true} \\ \text{false} & \text{false} & \text{true} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A inside B	$\begin{bmatrix} \text{true} & \text{false} & \text{false} \\ \text{true} & \text{false} & \text{false} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A contains B	$\begin{bmatrix} \text{true} & \text{true} & \text{true} \\ \text{false} & \text{false} & \text{true} \\ \text{false} & \text{false} & \text{true} \end{bmatrix}$	
A equals B	$\begin{bmatrix} \text{true} & \text{false} & \text{false} \\ \text{false} & \text{true} & \text{false} \\ \text{false} & \text{false} & \text{true} \end{bmatrix}$	

A meets B	$\begin{bmatrix} \text{false} & \text{false} & \text{true} \\ \text{false} & \text{true} & \text{true} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A covers B	$\begin{bmatrix} \text{true} & \text{true} & \text{true} \\ \text{false} & \text{true} & \text{true} \\ \text{false} & \text{false} & \text{true} \end{bmatrix}$	
A coveredBy B	$\begin{bmatrix} \text{true} & \text{false} & \text{false} \\ \text{true} & \text{true} & \text{false} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	
A overlaps B	$\begin{bmatrix} \text{true} & \text{true} & \text{true} \\ \text{true} & \text{true} & \text{true} \\ \text{true} & \text{true} & \text{true} \end{bmatrix}$	

DE-9-Intersection-Modell

- Dimensionally Extended 9 Intersection Model
- anstelle boolescher Werte Dimensionsangaben
- **d** gibt maximale Dimension an:
 - **-1**, falls $A \cap B = \emptyset$
 - **0**, falls $A \cap B$ nulldimensional/ **1**, falls eindimensional/ usw.
 - ***** falls $d(x) = \{-1,0,1,\dots\}$ „don't care“

Räumliche Datenbankabfragen



Basisanfragen:

- Punktabfragen
- Rechteckanfragen
- Regionsanfragen
- Abstandsanfragen
- Richtungsanfragen

Verbundanfragen:

- Intersection
- Difference
- usw.

Nächste-Nachbarn-Anfragen:

- Normale NNA
- K-nächste-NA
- Inkrementelle NNA

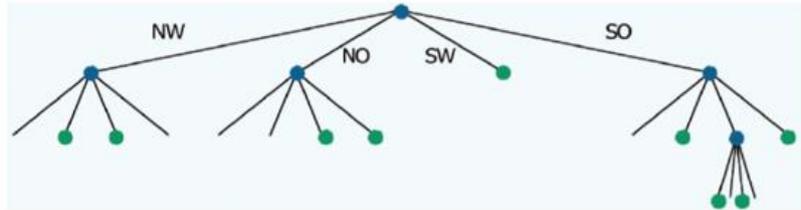
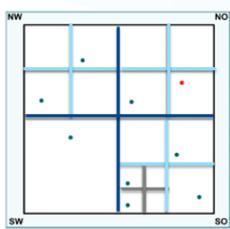


Indexierung von Geodaten

Quadtrees

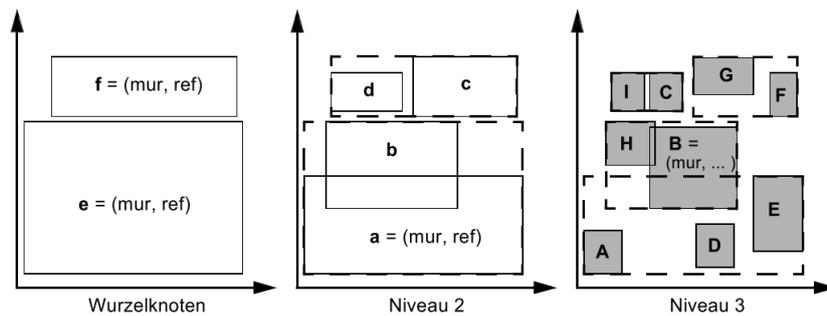
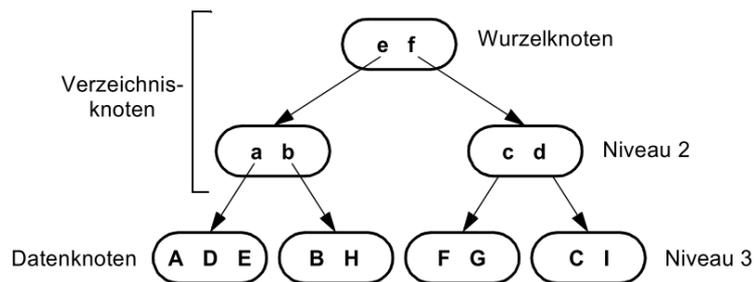
- Räumliche Datenstrukturen
- Prinzip: k-dimensionaler Datenraum wird rekursiv in 2^k gleich große Zellen („Kacheln“) unterteilt

- Varianten für Punkte, Linien und Flächen



R-Bäume

- organisieren k-dimensionale Rechtecke mithilfe überlappender Blockregionen
- Verzeichnisknoten umfasst Einträge der Form (mur, ref)
 - **mur** = minimal umgebendes Rechteck
 - **ref** = Verweis auf direkte Nachfahren



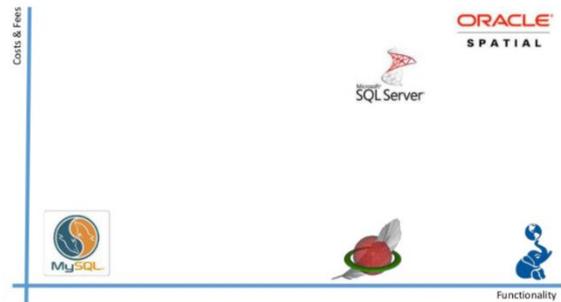
Geodatenbanksysteme

Kommerziell:

- Oracle Spatial (Oracle Locator)
- SAP SQL Anywhere
- IBM DB2 Spatial Extender
- Microsoft SQL Server

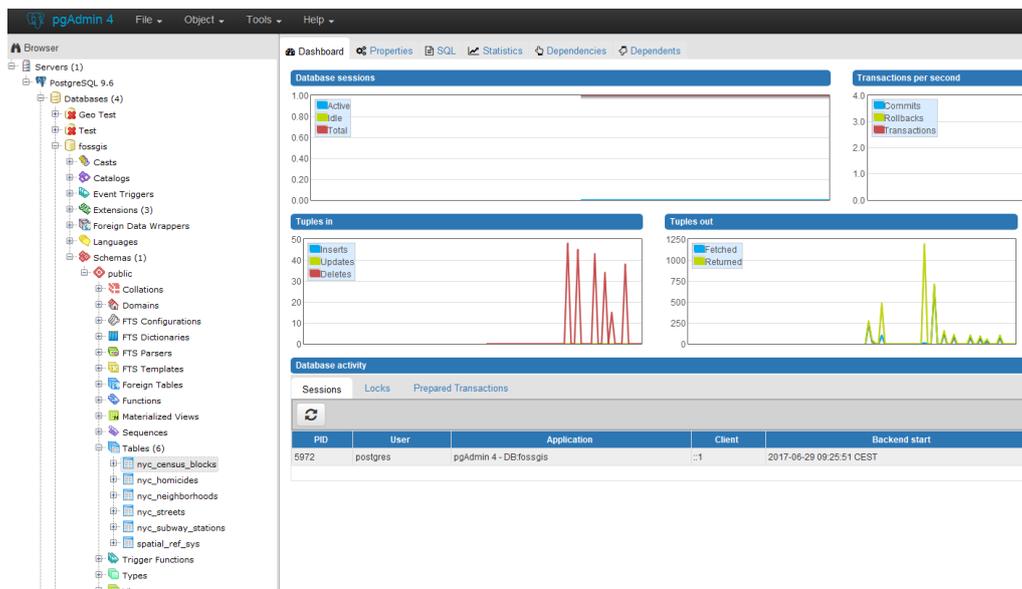
Open Source:

- MySQL (Spatial Extensions)
- Spatialite
- PostgreSQL (PostGIS)



Praxisteil PostGIS

Im Praxisteil des Referats wurde die PostgreSQL Erweiterung PostGIS vorgestellt und einige beispielhafte Datenbankabfragen durchgeführt. Die Verwaltung der Datenbank wurde mit dem Open Source Management Tool pgAdmin (Version 4) demonstriert.

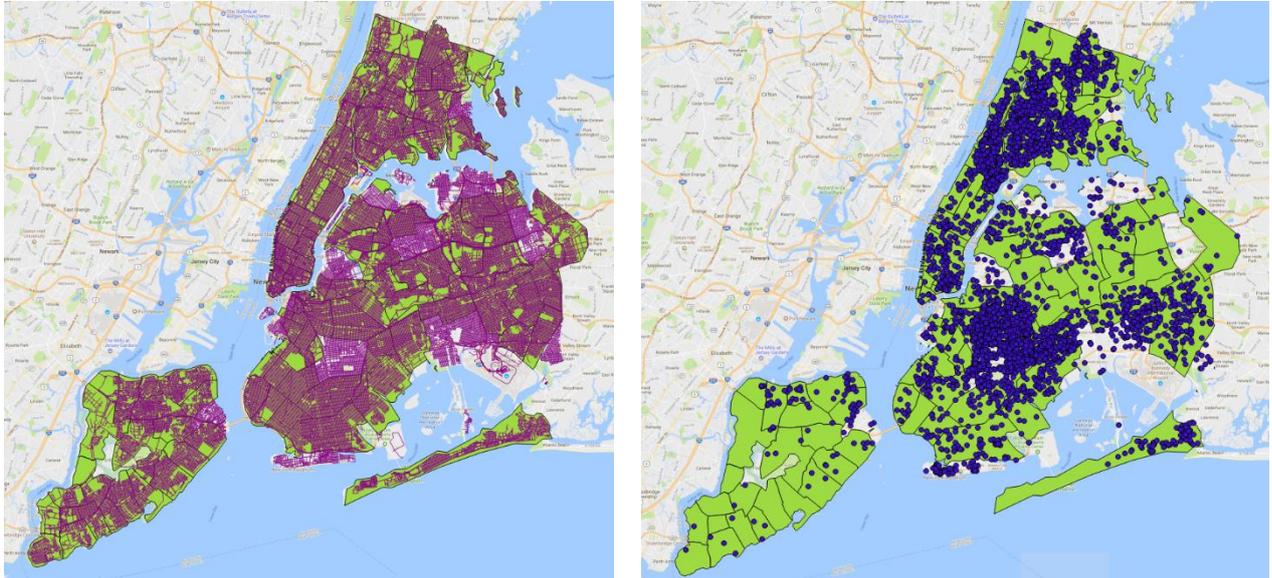


Als Beispieldaten wurden geografische (Straßen, Bezirke, Subbezirke, U-Bahn-Stationen) und soziodemografische Daten (Population, Mordstatistik) des US-Bundesstaats New York verwendet¹. Diese Daten wurden in Vorbereitung der Präsentation mittels des pgShapeLoaders in die PostgreSQL - Datenbank importiert. Das besondere Merkmal dieser Daten ist

¹ Diese können unter dem folgenden Link gedownloadet werden:
<http://files.boundlessgeo.com/workshopmaterials/postgis-workshop-201401.zip>

die Spalte Geometry, welche eine Repräsentation der Objekte im WKB (Well-known-Binary) Format beinhaltet.

Die grafische Darstellung der Datensätze erfolgte mit der Open-Source Software QGIS, welcher zu Anzeige, Bearbeitung und Analyse räumlicher Daten dient.



Die Demonstration beinhaltete zusätzlich zur Erläuterung der Darstellung einzelner Datenlayer eine beispielhafte Analyse der Daten. Dazu wurde eine konvexe Hülle um sämtliche Bezirke gebildet und die darin enthaltenen Mordfälle (welche als Punkte repräsentiert werden) mittels Analysewerkzeug gezählt.

Abschließend wurden in pgAdmin verschiedene geografische Abfragen auf die Beispieldaten angewendet, welche nachfolgend auszugshaft aufgelistet sind:

1. Geometrieausgabe im WKT (Well-known-Text) Format:

```
SELECT ST_AsText(geom) FROM nyc_subway_stations;
```

2. Arealgröße der Bronx (in km²):

```
SELECT SUM(ST_Area(geom) / 1000000) FROM nyc_neighborhoods WHERE boroname = 'The Bronx';
```

3. Länge aller Straßen New Yorks (in km):

```
SELECT Sum(ST_Length(geom)) / 1000 FROM nyc_streets;
```

4. Alle Straßen im Umkreis von 100m von der U-Bahn-Haltestelle Broad St.:

```
SELECT name FROM nyc_streets WHERE ST_DWithin(geom, ST_GeomFromText('POINT(583571 4506714)', 26918), 100 );
```

5. Straßen, die Atlantic Common kreuzen:

```
SELECT name FROM nyc_streets WHERE ST_DWithin(geom, ST_GeomFromText('LINESTRING(586782 4504202, 586864 4504216)', 26918), 0.1);
```

Quellen und weiterführende Links

- Brinkhoff, Thomas: Geodatenbanksysteme in Theorie und Praxis, Berlin 2008
- Kudraß, Thomas (Hg.): Taschenbuch Datenbanken, München 2015
- Störl, Uta: Geodatenbanksysteme, URL: <https://www.fbi.h-da.de/fileadmin/personal/u.stoerl/DBAkt-WS1415/Vorlesung/DBAkt-WS1415-GeoDB-Kap1.pdf>
- Güting, Ralf Hartmut: An Introduction to Spatial Database Systems, Hagen 1994
- Jung, Christopher: Vergleich von Quadtree, kd-tree und r-tree für statische und dynamische Geodaten (Masterthesis), Dortmund 2011
- Dallüge, Uwe: Introduction to PostGIS, URL: <http://workshops.boundlessgeo.com/postgis-intro/>