# **Allgemeines**

#### Modul

• Methoden der Geographie III: Geogr. Informationssysteme (GIS)

#### Semester

SS18

## Übung

- 5 Übungsblätter (50%)
- Abschluffprojekt (50%)
  - 10% Form
  - 80% Inhalt
  - 10% Karte

#### **GIS Software**

### **QGIS**

https://qgis.org/de/site/

### **ArcGIS**

• https://www.arcgis.com/features/index.html

### Literatur & weiterführende Materialien

## **Geographic Information System Basics** (kostenlos)

• https://2012books.lardbucket.org/books/geographic-information-system-basics/index.html

#### **ArcGIS - Hilfe**

- Durchsuche die Hilfe nach Begriffe (Fachwörter, Konzepte, etc.)
- https://desktop.arcgis.com/de/arcmap/

### **Tools**

Siehe: http://kollektive-geographie-heidelberg.de/#/geographie-ressourcen/tools

# Einführung

## Grundbegriffe

#### GIS

- Geographisches Informationssystem
- a computerized system that can store, retrieve, manipulate, visualize and analyse spatial data
- ermöglicht die Integration unterschiedlicher Arten von Daten in einem System

## Informationssystem

- Erfassung
- Speicherung
- Aktualisierung
- Verarbeitung
- Wiedergabe

### **Datenverarbeitung**

- Erfassung
- Verwaltung
- Analyse
- Modellierung
- Visualisierung

### Geodaten (Spatial Data)

- Geographische Daten
- Informationen über die Lage und Form (Geometriedaten) einer Erscheinung (Objekt) auf der Erdoberfläche und über die (nicht geometrischen) Eigenschaften (Attributdaten) dieser Erscheinung (Spektrum).
- Geodaten beschreiben folgende Merkmale von Geoobjekten:
  - Geometrie
  - Topologie
  - Thematik
  - Dynamik

### Geoobjekt (Features)

• Ein auf der Erde vorhandenes Objekt, das mittels Geodaten eindeutig referenzierbar ist (Wikipedia).

#### **Attributdaten**

- Eigenschaften eines Geoobjekte
- Daten ohne spezifischen Raumbezug (Spektrum)
  - Bei Vektordaten in der Regel in konventionellen Datenbanken gespeichert
  - Bei Rasterdaten entspricht der Wert des Attributs dem gespeicherten Wert der Rasterzelle

### Layerprinzip

- Verschiedene Objekt-klassen oder Variablen in Schichtmodell (Layer) miteinander verknüpft
- jede Geometrietyp (oder Attributklasse) wird in eigener Informationsschicht erfasst

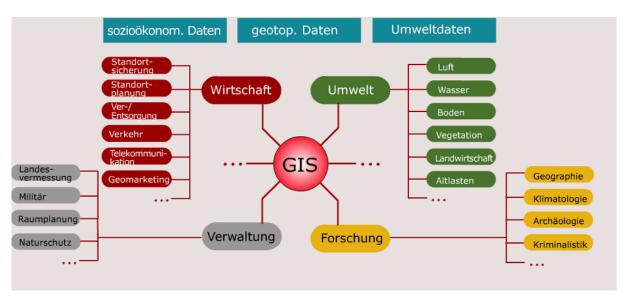
#### **EVAP**

- 1. Erfassung
- 2. Verwaltung
- 3. Analyse
- 4. Präsentation

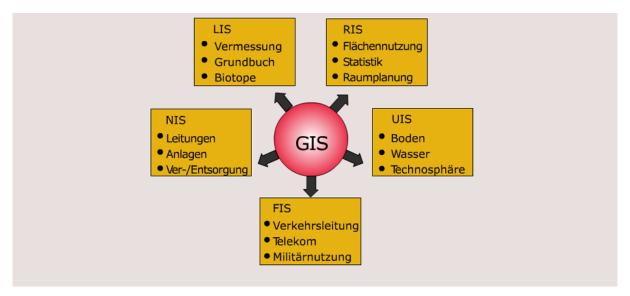
## Raumbezogene Datenanalyse

- 1. Geometrische-topologische Verfahren
- 2. Statistische Analysen und Mengenoperatoren
- 3. Modellierung und Simulation

## **Anwendungsfelder und GIS-Typen**



Quelle: http://www.geoinformation.net/lernmodule/folien/Lernmodul\_03/Lerneinheit01/index. html



Quelle: http://www.geoinformation.net/lernmodule/folien/Lernmodul\_03/Lerneinheit01/index. html

- LIS = Land(schafts)informationssystem (z.B. zur Naturraumausstattung)
- RIS = Rauminformationssystem (z.B. für Regionalplanung)
- KIS = Kommunales Informationssystem (für Planung u. Verwaltung)
- UIS = Umweltinformationssystem (z.B. Umweltüberwachung)
- NIS = Netzinformationssystem (z.B. Kanal-IS, Straßen-IS)
- FIS = Fachinformationssystem (z.B. Boden-IS, ...)

#### Raummodelle

In order to visualize natural phenomena, one must first determine how to best represent geographic space. Data models are a set of rules and/or constructs used to describe and represent aspects of the real world in a computer. Two primary data models are available to complete this task: raster data models and vector data models.

#### Ziel von GIS

 rechnergestützte raumbezogene Analysen mit Geodaten erforderlich: digitales Modell der "Wirklichkeit"

### Modelle

- 1. Vektormodell
  - · Welt als leerer Raum, der mit diskreten Objekten (Entitäten) angefüllt ist
  - Vector data models use points and their associated X, Y coordinate pairs to represent the vertices of spatial features,

hree fundamental vector types exist in geographic information systems (GISs): points, lines, and polygons

### 2. Rastermodell

- Welt als Reihe von Variablen, die an jeder Stelle einen Wert annehmen
- The raster data model consists of rows and columns of equally sized pixels interconnected to form a planar surface.

# **Analysemethoden**

Dieses Kapitel ist unabhängig von den Datenmodellen (Rasterdatenmodell und Vektrodatenmodell). Siehe in den jeweiligen Kapiteln für Analysemethoden von Rasterdaten oder Vektordaten.

### **Abfragevarianten**

#### **Thematische Abfragen**

• Selektiert die Objekte, deren Eigenschaften (Attribute) die gestellten Bedingungen erfüllen.

#### **Geometrische Abfragen**

• Selektiert die Objekte, die bestimmte räumliche Bedingungen erfüllen.

### **Topologische Abfragen**

• Selektiert die Objekte, welche die gestellten Bedingungen bezüglich den räumlichen Beziehungen zwischen den Objekten erfüllen.

### Kategorien räumlicher Analyse

- Location: Was ist Wo?
  - Welche räumlichen (Fläche, Umrisslänge, Mittelpunkt etc.) und nicht-räumlichen Attribute besitzt ein Objekt (z.B. Einwohnerzahl einer Gemeinde).
- Condition: Wo ist Was?
  - Auffinden von räumlichen Objekten mit bestimmten inhaltlichen Voraussetzungen (z.B. Gebäudekataster: Wo befinden sich sanierungsbedürftige Altbauten? ).
- Trend: Zeitl. Entwicklung
  - Untersuchung von Entwicklungen mit Flächenverschneidungen (z.B. Ausweitung von besiedelter Fläche zu zwei verschiedenen Zeitpunkten).
- Pattern: relevante Muster

- Untersuchung von raumbezogenen Korrelationen mit Flächenverschneidungen (z.B. Häufung bestimmter Erkrankungen in Abhängigkeit von emittierenden Betrieben).

## Routing:

- Netzwerkanalysen z.B. bei der Verkehrsplanung (z.B. Auffinden des kürzesten Wegs zwischen zwei Städten unter Benutzung ausschließlich von Bundesstrassen?).
- Modelling: Prozesse ab(nach)bilden
  - Gemeint ist die Modellierung von Raumszenarien (z.B. Simulation von Hochwasserereignissen).

### **Auswahloperatoren**

## **Vergleichende Operatoren**

- = (gleicht)
- > (größer)
- < (kleiner)</li>
- <> (ungleich)

### **Arithmetische Operatoren**

- (Multiplikation)
- / (Division)
- (Addition)
- (Subtraktion)
- exp (Exponent)
- % (Modulo)

## **Logische Operatoren**

- AND ☑ (Schnittmenge)
- OR ☑ (Vereinigungsmenge)
- XOR (ausschließende (exklusive) Vereinigungsmenge)
- NOT ¬ (Negation)

## Verschneidung

### Verschneidung

 Gruppe grundlegender GIS-Funktionen, die ein digitales Zusammenführen (Kombination) von Geo- und Attributdaten mehrerer Themenebenen oder Objektklassen ermöglicht (Geoinformatik - Uni Rostock)

### **Boolesche Verschneidung**

Verschneidung auf binärer Informationsebene (wahr/falsch).

Operator	Operatorname	Beschreibung
AND	Schnittmenge	Ergibt "wahr" für alle Gebiete, die sowohl das erste als auch das zweite Kriterium
OR	Vereinigung	Ergibt "wahr" für alle Gebiete, die entweder das erste oder das zweite Kriterium
XOR	exklusive Vereinigung	Ergibt "wahr" für alle Gebiete, die entweder das erste oder das zweite Kriterium
NOT	Negation	Ergibt "wahr" für alle Gebiete, die das erste Kriterium erfüllen, nicht aber das zw

Quelle: http://gitta.info/Suitability/de/html/BoolOverlay\_learningObject2.html

### **Geometrische Verschneidung**

### Operationen der Mengenlehre

- Vereinigung (Union) = OR
- Schnitt (Intersection) = AND
- Symmetrische Differenz (Symmetrical difference) = XOR
- Mengen Differenz (Identity)

#### Arten der Verschneidung

- Punkt-Fläche
- Linie-Fläche
- Fläche-Fläche

## Rasterdaten

#### Rasterdatenmodell

• Ein Datenmodell, in dem Bildinhalte (z.B. Fotos) oder räumliche Objekte als (quadratische) Rasterzellen in einer zweidimensionalen Datenmatrix abgebildet werden. (Spektrum)

- Ein Gitter (Grid) aus Zeilen und Spalten (Wikipedia)
- Eine Menge von Zellen (Pixel)
- Zellen haben eine homogene Größe
- Einzelnen Zellen werden Werte zugeordnet, durch die der in der Zelle abgebildete Raum beschrieben wird
- Werte von Rasterdaten
  - Floating Point-Grid
    - \* für kontinuierliche Daten
    - \* Wert der Zelle = Attribut
  - Interger-Grid
    - \* für diskrete Daten
    - \* Wert der Zelle ist einem bestimmten Attribut zugeordnet (Value Attribute Tabel)
- Geeignet bei kontinuierlicher Phänomenen
- Weniger geeignet für diskrete Phänomene

Vorteil	Nachteile
Technologisch schnelle und günstige Erfassung von Daten	großer Speicherplatzb
Einfache Datenstruktur und resultierend daraus einfache Abfrage- und Analysealgorithmen	Maßstabsabhängig (Ei
	Probleme in der Besch
	Es gibt keine Informati

## **Typen von Rasterdaten**

- Bildraster
  - Luft-/Satellitenbilder
- Thematische Raster
  - Landnutzung
- Kontinuierliches Raster
  - Höheninformation

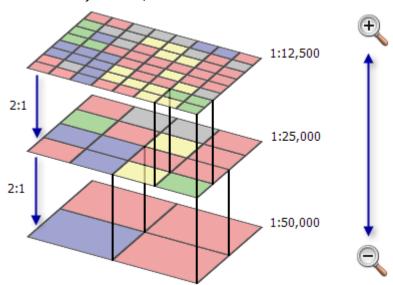
## Rasterformate

- ASCII Grid
  - Textuell (Menschenlesbar)
  - Rasterzellen als Matrix

- GeoTIFF
  - Binärformat
  - Attributdaten in Hauptdatei integriert
- GeoJPEG
  - Binärformat
  - Attributdaten in gesonderter Datei
- Geopackage und Geodatabase

## **Raster Pyramiden**

Pyramiden werden zur Verbesserung der Performance verwendet. Sie sind eine reduzierte Version des ursprünglichen Raster-Datasets und können viele reduzierte Layer enthalten. Für jeden Layer der Pyramide wird ein Resampling im Verhältnis 2:1 durchgeführt. Unten findet sich ein Beispiel für zwei Ebenen von Pyramiden, die für ein Raster-Dataset erstellt wurden (ESRI - ArcMap):



Pyramiden können die Anzeige von Raster-Daten beschleunigen, da nur die für die Anzeige erforderlichen Daten in der vorgegebenen Auflösung abgerufen werden.

## Analysemethoden: Karten Algebra (Map Algebra)

## **Definition**

- Methodengebäude und "Tool-Box" zur Kombination von zwei oder mehr Rasterdatensätzen oder Layern (Rasterdaten).
- Map algebra wendet mathematische Operationen auf komplette Rasterdatensätze an, wobei die jeweilige Operation für jede einzelne Rasterzelle in jedem Datensatz durchgeführt wird.

- Hierdurch lassen sich die als Zahlenmatrizen gespeicherten Informationsebenen miteinander verrechnen. (Spektrum)

### **Mathematische Operationen**

- Arithmetisch
- Logarithmisch (Log, ...)
- Trigonometrisch (Sin, ...)
- Power (Wurzel, ...)

### Typen der Operatoren

- Lokal
  - Punktoperationen (Pixel)
- Fokal
  - Nachbarschaftsoperationen (Distanzoperationen)
  - direkten Nachbarzellen werden in die Berechnung mit einbezogen.
- Zonal
  - Gebiete gleicher Klassenzugehörigkeit
- Global
  - Thematisch/komplette Ebene

#### **Nachbarschaften**

#### **Definition**

- größere Zonen auf Basis von Distanzen und/oder Richtungen.
- N4-Nachbarschaft
  - Lateraler Nachbar oder Diagonaler Nachbar
- N8-Nachbarschaft
  - Lateraler Nachbar *und* Diagonaler Nachbar

# **Beispiel: Minesweeper**

• als klassisches Beispiel einer 3 x 3- Zellen Nachbarschaftsanalyse: Jeder Zellwert (Nummer) gibt die Anzahl der Bomben innerhalb einer 3-Zellen-Nachbarschaft an.

## **Analyse Prinzip**

#### Verschneidung

*Gegeben* sind zwei Raster. Ein thematisches Raster mit Landnutzungsinformationen (Wald, Wasser, Agrar, etc.)) und ein kontinuierliches Raster mit Vegetationsindex.

Gesucht sind die Agrarflächen mit einem bestimmten Vegetationsindex.

- Landnutzungsraster
  - Setze Pixel mit dem Wert Agrar auf 1
  - Setze alle anderem Pixel auf den Wert 0
- Vegetationsindexraster
  - Setze alle Pixel mit dem gewünschten Vegetationsindex auf 1
  - Setze alle anderen Pixel auf den Wert 0
- Verschneidung
  - Verschneide beide Raster (logisches AND)

#### Vektordaten

#### Vektordatenmodell

#### Vektordaten

• Daten entsprechend des Vektordatenmodells zur Abbildung von Punkten, Linien, Flächen und Körpern auf der Basis von Koordinaten. Neben den Koordinaten als Träger geometrischer Informationen können, je nach Datenformat, weitere objektbezogene Informationen in Form von topologischen, graphischen oder Sachdaten in einem Vektordatensatz enthalten sein (Spektrum).

#### Vektordatenmodell

- Vektordaten beschreiben raumbezogene Objekte anhand von Punkten
- Geeignet für diskrete Phänomene
- Weniger geeignet für kontinuierliche Phänomene
- Im Vektordatenmodell wird ein *Punktobjekt* als einzelnes Koordinatenpaar (x/y) erfasst
  - Linieobjekte werden über Listen von Punkten repräsentiert
  - Polygone sind geschlossenen Linienzügen, sowie den Polygonzentroiden (Zentroid)
- Jedes Objekt kann neben seiner Geometrie auch Eigenschaften besitzen (Attributdaten).
  - Im Gegensatz zum Rastermodell sind diese sog. Attributdaten mit dem Objekt selbst verknüpft.
- Im Vektormodell unterscheidet man zwischen topologischen und nichttopologischen Datenstrukturen.
  - Nichttopologische Datenstrukturen (Spaghetti-Datenstrukturen) bilden lediglich die Lage und Form eines Objektes ab (Geometriedaten), enthalten aber keine Informationen über Nachbarschaftsbeziehungen. (nur Geometrie, keine Topologie)

- Topologische Vektordaten enthalten hingegen zusätzlich Informationen über die räumlichen Beziehungen der Objekte. (Geometrie und Topologie)

Quelle: (Spektrum)

Vorteil	Nachteile		
Scharfe Repräsentation der Geometrie	Komplexität: komplexe Datenstruktur und resultierend d		
Beliebige unregelmäßige Formen und Rauminhalte	Schwierig zu bearbeiten		
Kompaktheit: Geringerer Speicherplatzbedarf			
Maßstabsunabhängig (keine Einschränkungen bei Zoom)			

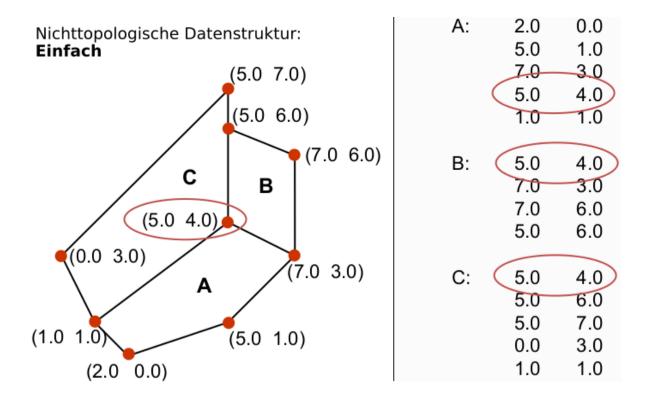
The final advantage of vector data is that topology is inherent in the vector model. This topological information results in simplified spatial analysis (e.g., error detection, network analysis, proximity analysis, and spatial transformation) when using a vector model (geographic information system basics).

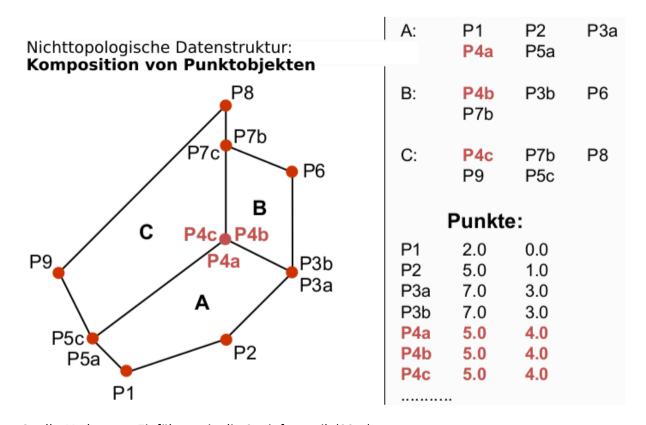
## Nichttopologische Vektordatenstrukturen (Spaghetti Modell)

- Reihung von Koordinatenpaaren
  - Punkte (XY-Koordinaten)
  - Linien (Reihe von Punkten, deren erster und letzter ungleiche Koordinaten besitzen)
  - Polygonen (Reihe von Punkten, deren erste und letzte gleiche Koordinate besitzen)

#### **Vor- und Nachtteile**

- Vorteile
  - einfache Datenstrukture
- Nachteile
  - Topologie nur implizit
  - Fehleranfällig durch Redundanz





Quelle: Vorlesung - Einführung in die Geoinformatik (SS18)

## Topologische Vektordatenstrukturen (Knoten- und Kantenstruktur)

Speichert topologischen Beziehungen explizit

- · Adjazenz von Knoten und Kanten
- Adjazenz von Kanten und Maschen
- Adjazenz von Kanten und Kanten
- Adjazenz von Maschen und Maschen

Maschen = Polygone

Adjazenz = Berühren, Aneinandergrenzen (Graphentheorie)

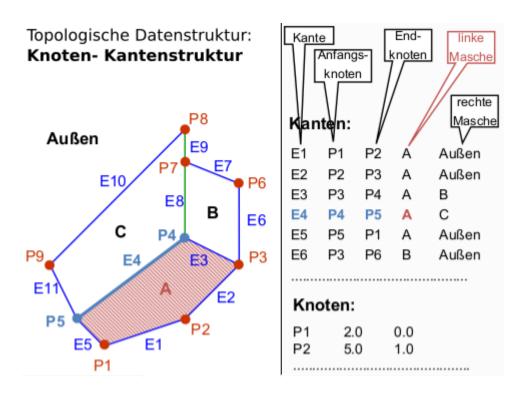
Siehe auch Graphen und Netzwerke

#### **Knoten- und Kantenstruktur**

- Knoten (XY-Koordinaten) werden gespeichert
- Kanten (Knotenpaare und angrenzende Maschen) werden separat von Knoten gespeichert

### **Vor- und Nachtteile**

- Vorteile
  - Geometrie ist redundanzfrei
  - Topologie ist explizit
  - Bei Änderungen können Fehler leichter vermieden werden
- Nachteil
  - größere Rechenkapazität im vgl. zu zu nichttopologischen Datenstrukuturen wird gebracht



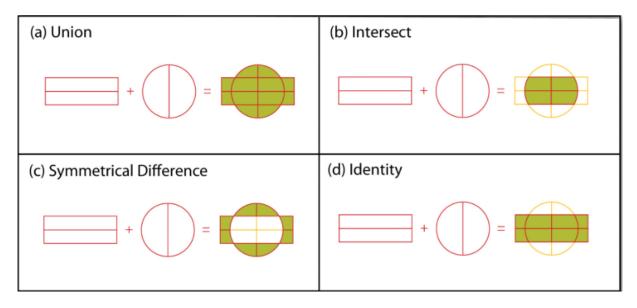
Quelle: Vorlesung - Einführung in die Geoinformatik (SS18)

## Analysemethoden

- Geometrisch
  - Manipulation und Nutzung der Form, Lage und Größe der Vektorfeatures
- Topologisch
  - Analyse der gegenseitigen Lagebeziehungen der Vektorfeatures

## **Geometrische Verschneidung (Overlay Operations)**

- Vereinigung (Union) = OR
- Überschneidung (Intersection) = AND
- Symmetrische Differenz (Symmetrical difference) = XOR
- Mengen Differenz (Identity)



Quelle: https://2012books.lardbucket.org/books/geographic-information-system-basics/index.html

### Geoprozessierung

Daten werden verändert, aber keine neuen "Informationen" geschaffen

- Auflösen (Dissolve)
  - Kombiniert mehrere Polygonobjekte zu einem Polygonobjekt, basierend auf einem bestimmten Attributen
- Auswählen (Select)
  - Erstellung eines Ausgabelayers basierend auf einer benutzerdefinierten Abfrage, welche bestimmte Objekte des Eingabelayers auswählt. (Siehe oben: Auswahloperatoren)
- Merge
  - Verbindet gleiche Objekte aus unterschiedlichen Eingabelayern in einem neuen Ausgabelayer
- Append
  - Zusammenführen von Layer mit existierenden Überschneidungsbereichen
- Differenz (Clip)
  - Extrahiert die Objekte eines Eingabelayers (Punkt, Linie oder Polygon), welche in die räumliches Ausdehnung (spatial extent) ein Clip Layers fällt
- Erase
  - Gegenteil von Clip
  - Behält die Objekte eines Eingabelayers (Punkt, Linie oder Polygon), welche außerhalb der räumliches Ausdehnung (spatial extent) ein Clip Layers sind
- Split

- Teilt einen Eingabelayer in mehrere Ausgabelayer (räumlich unterschiedliche Teilbereiche) auf

### **Puffer (Buffer)**

## Puffer (Buffer)

• Das Erzeugen eines Polygonlayers, welcher eine Zone (oder mehrerer Zonen) mit einer bestimmten Breite um ein Eingabe Punkt-, Linie- oder Polygonobjekt

### **Eingabe**

- Punkt-, Linie- oder Polygonobjekt
- Breite

### **Ausgabe**

• Polygonlayer mit Zonen

#### **Typen**

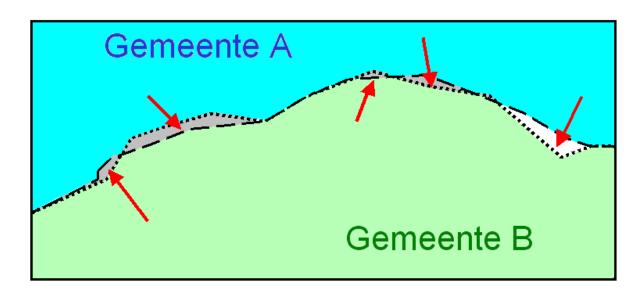
- · Constant width buffers
- · Variable width buffers
  - Pufferbreite wird vorher für jedes Objekt festgelegt
  - Abhängig von Attributen des jeweiligen Objektes

### **Multiple ring buffers**

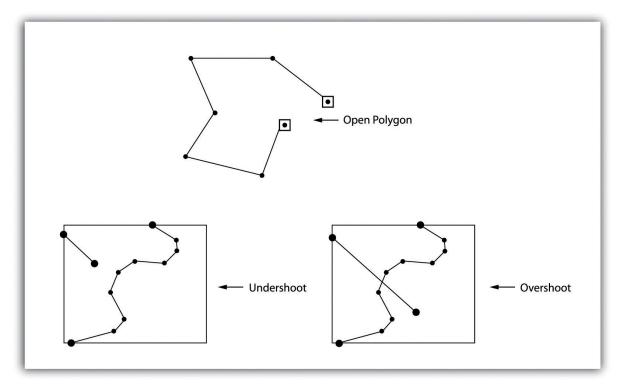
• Serie von konzentrischen Pufferzonen (wie eine Zielscheibe) um ein Objekt.

## **Topologische Fehler**

**Sliver Polygone** sind fehlerhaft entstandene Restflächen, die beim räumlichen Verschneiden verschiedener Datenebenen entstehen (Wikipedia). Zwei Kanten von Polygonen treffen nicht genau aufeinander.



Weitere Fehler sind Offene Polygone, Undershots und Overshots.



### **Work in Progress**

#### Nächste-Nachbar-Analyse

k-Nächste-Nachbar-Analyse (euklidische Distanz)

#### Geometrisch: Geometrien vereinfachen

Problemstellung: Ihre Geometrien sind für einen gewünschten kartographischen Darstellungsmaßstab zu detailliert. Lösung: Vereinfachung über (z.B.) den Douglas-Peucker-Algorithmus zur geometrischen Vereinfachung.

Änderung des Darstellungsmaßstabes und Einsparung von Speicherplatz (weniger Punkte).

# **Datenmodellierung**

#### **Datenmodell**

- Abbildung der realen Welt in abstrahierter, vereinfachter, strukturierter Weise
- Ein (relationales) Datenmodell wird durch die Abstraktion von einzelnen Objekten (Entitäten), [deren Beziehungen zueinander (Relationen)] und deren Eigenschaften (Attribute) gebildet (http://geoinformatik.lehrewelt.de)

#### Phasen der Datenmodellierung

- 1. Konzeptionelles Datenmodell (Datenbankschema)
  - Bestimmung und Benennung der wichtigsten Objekte (Entitäten)
  - Bestimmung der wichtigsten Beziehungen zwischen Objekten (Relationen)
- 2. Logisches Datenmodell (Datenbankschema)
  - Bestimmung der relevanten Eigenschaften von Objekten (Attribute)
  - Beschreibung des Verhaltens von Objekten (Methoden)
- 3. Physisches Datenmodell (Datenbankschema)
  - Definition der Datentypen für Objekte und deren Eigenschaften
  - Definition wie Beziehungen implementiert werden sollen (Kardinalitäten)
  - Beschreibung der Operationen und deren Parameter (Methodendefinition/-deklaration)

### **Entity-Relationship Modell (ERM)**

Das Entity-Relationship-Modell dient dazu, im Rahmen der semantischen Datenmodellierung einen in einem gegebenen Kontext, relevanten Ausschnitt der realen Welt zu beschreiben.

Das ERM besteht aus einer Grafik (ER-Diagramm) und einer Beschreibung der darin verwendeten Elemente, wobei deren Bedeutung (Semantik) und ihre Struktur dargestellt wird (Wikipedia)

### **Begriffe**

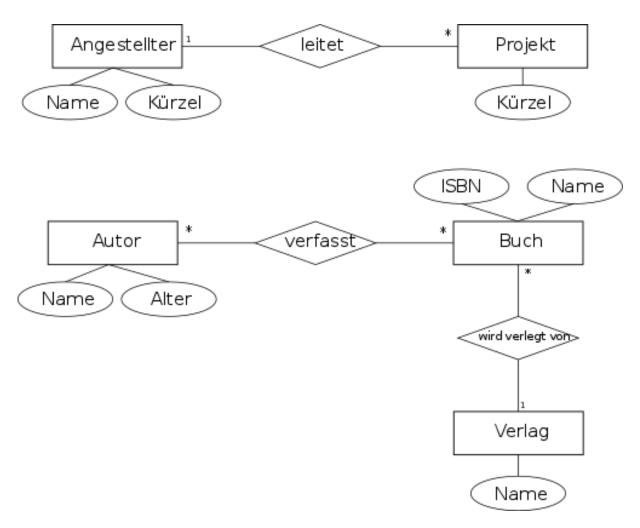
Objekte und Gegebenheiten der realen Welt:

- Entität (Entity)
  - individuell identifizierbares Objekt der Wirklichkeit
  - z. B. der Angestellte Müller oder das Projekt 3232
- Beziehung (Relationship)
  - Beziehung zwischen zwei oder mehreren Entitäten
  - z. B. Angestellter Müller leitet Projekt 3232
- Eigenschaft (Attribute)
  - Eigenschaft von einer Entität
  - z. B. das Eintrittsdatum des Angestellten Müller.

Im Rahmen der Modellierung werden geichartige Typen gebildet:

### Entitätstyp

- Typisierung gleichartiger Entitäten
- z. B. Angestellter, Projekt, Buch, Autor, Verlag
- Beziehungstyp (Relationship-Typ)
  - Typisierung gleichartiger Beziehungen
  - z. B. Angestellter leitet Projekt Attribut
  - Typisierung gleichartiger Eigenschaften
  - z. B. Nachname, Vorname und Eintrittsdatum im Entitätstyp Angestellter.



#### Kardinalitäten

- Die Kardinalität legt (auf der Ebene Beziehungstyp) für jeden der beteiligten Entitätstypen fest, an wie vielen konkreten Beziehungen (dieses Typs) seine Entitäten beteiligt sein können oder müssen (Wikipedia).
- Es existieren unterschiedliche Notationsformen
- Chen-Notation
  - 1:1
  - 1:n
  - m:n

Quelle: https://de.wikipedia.org/wiki/Entity-Relationship-Modell

### **Vor- und Nachteile**

#### Vorteile

- · einfach und intuitiv
- einfach abzubilden auf relationales Modell
- weite Verbreitung

#### Nachteile

- geringe Ausdrucksfähigkeit
- formal nicht eindeutig
- keine Spezifikation von Verhaltensaspekten

## Spezialisierung und Generalisierung (is-a-Beziehung)

Bei der **Spezialisierung** wird ein Entitätstyp als Teilmenge eines anderen Entitätstyps erkannt und deklariert, wobei sich die spezialisierte Entitätsmenge durch besondere Eigenschaften (nur für sie geltende Attribute und/oder Beziehungen) gegenüber der übergeordneten, generalisierten Menge auszeichnet (Wikipedia).

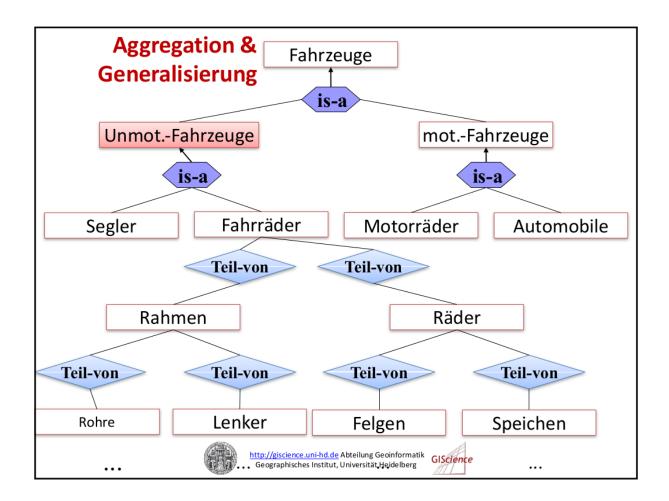
### Aggregation und Zerlegung (is-part-of-Beziehung)

Werden mehrere Einzelobjekte (z. B. Person und Hotel) zu einem eigenständigen Einzelobjekt (z. B. Reservierung) zusammengefasst, dann spricht man von **Aggregation**. Dabei wird das übergeordnet eigenständige Ganze Aggregat genannt; die Teile, aus denen es sich zusammensetzt, heißen Komponenten. Aggregat und Komponenten werden als Entitätstyp deklariert (Wikipedia)

### Aggregation vs. Generalisierung

Aggregation und Generalisierung bilden Hierarchien aber:

- 1. Aggregation bezieht Objekte aufeinander
- 2. Generalisierung bezieht Klassen aufeinander



## **Datenbanken**

## **Begriffe**

## Datenbank (DB)

- Strukturierte Sammlung von Daten
- Menge der zu verwaltenden Daten

### **Datenbankmanagementsystem (DBMS)**

- Verwaltungsoftware
- organisiert intern die strukturierte Speicherung der Daten und kontrolliert alle lesenden und schreibenden Zugriffe auf die Datenbank (Wikipedia)

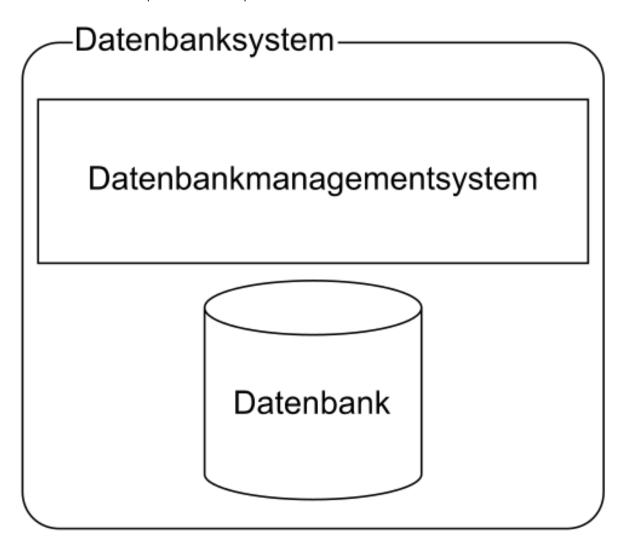
### Datenbanksystem (DBS)

• DB + DBMS

• Zur Abfrage und Verwaltung der Daten bietet ein Datenbanksystem eine Datenbanksprache an (Wikipedia)

#### **Datenbankmodell**

- Das Datenbankmodell bestimmt, in welcher Struktur Daten in einem Datenbanksystem gespeichert werden (Wikipedia)
- Modelle
  - Hierarchisch (One-to-many association across levels; z.Bsp. Stammbaum)
  - Netzwerk (Many-to-many)
  - Relational (Tabellenbasiert)



## Datenunabhängigkeit

• Der Grad den ein Benutzer auf die Daten eines Datenbanksystems zugreifen kann, ohne Details der systemtechnischen Realisierung der Datenspeicherung und des Datenzugriffs zu kennen

(Wikipedia)

- Benutzer interagiert nicht direkt mit Daten, sondern mit Repräsentation der Daten über standardisierte Datenbanksprache
- Physische Datenunabhängigkeit
  - Änderungen auf der physischen Ebene haben keinen Einfluss auf die logische Struktur
- Logische Datenunabhängigkeit
  - Änderungen auf der logischen Ebene haben keinen Einfluss auf Anwendungen

#### **Funktionen eines DBMS**

- Datenbanksprache
  - Funktionen
    - \* Datenabfrage und -manipulation (DML)
    - \* Verwaltung der Datenbank und Definition der Datenstrukturen (DDL)
  - Bei relationalen DBMS sind diese Kategorien in einer Sprache (SQL) vereint
- Verwaltung der Metadaten
- Datensicherheit
  - Schutz gegen Datenverlust (Sicherung/ Backup)
  - Schutz gegen unerlaubten Zugriff (Zugriffsrechte)
- Datenintegrität
  - Die Integrität der Daten kann durch Constraints sichergestellt werden
  - Regeln im Managementsystem, die beschreiben, wie Daten verändert werden dürfen
- · Mehrbenutzerfähigkeit
  - Zugriffsberechtigungen (Permissions)
- Transaktionen
- Abfrageoptimierung

Quelle: https://de.wikipedia.org/wiki/Datenbank

### **Transaktion**

#### **Transaktionen**

- Ermöglichung des Mehrbenutzerbetriebs durch das Transaktionskonzept
- Daten gegen Race Conditions durch den parallelen Zugriff mehrerer Benutzer schützten

## **Eigenschaften von Transaktionen (ACID)** (Wikipedia)

• atomicity = Atomarität (Abgeschlossenheit)

- Datenbank-Operationen, die entweder ganz oder gar nicht ausgeführt werden (Alles-odernichts-Eigenschaft)
- consistency = Konsistenzerhaltung
  - Nach Beendigung einer Transaktion wird ein konsistenter Datenbankzustand hinterlassen
  - Alle im Datenbankschema definierten Integritätsbedingungen vor dem Abschluss der Transaktion werden überprüft
- isolation = Isolation (Abgrenzung)
  - Verhindert, dass sich nebenläufig in Ausführung befindliche Transaktionen gegenseitig beeinflussen
  - Realisiert durch Zugriffssperren
- durability = Dauerhaftigkeit
  - Daten sind nach dem erfolgreichen Abschluss einer Transaktion garantiert dauerhaft in der Datenbank gespeichert

## Relationale Datenbanken

### Konzept

Eine **relationale Datenbank** beruht auf einem **tabellenbasierten relationalen Datenbankmodell** (Wikipedia):

- Eine **relationale Datenbank** kann man sich als eine Sammlung von **Tabellen (Relationen)** vorstellen.
- Jede Zeile (Tupel) in einer Tabelle ist ein Datensatz (Record).
- Jede \*\*Spalte (Attribute) \*\* besteht aus einer Reihe von **Attributwerten (Attribute)** 
  - Alle Attributwerte eines Attributes (Spalte) ist von einem bestimmten Datentyp
- Die Tabellen sind durch Beziehungen (relationships) miteinander verknüpft
  - Dadurch ist es möglich bei Abfragen, Daten aus mehreren Tabellen zu kombinieren
- Ein Datensatz muss eindeutig identifizierbar sein. Das geschieht über einen oder mehrere Schlüssel.

relationales Modell	informeller Begriff	Erklärung
Relation, Entitätstyp, Entitätsklasse	(table)	Eine Tabelle in einer Datenbank
Tupel, Entität (tuple, entity)	Zeile (row)	Ein horizontaler Datensatz einer Tabelle in der Datenbanl
Beziehung (relationship)		Beziehung einzelner Tupel zueinander
Kardinalität (cardinality)		Mengenangabe zur Beziehung einzelner Tupel (z. B. 1:1, 1:

relationales Modell	informeller Begriff	Erklärung
Attribut (attribute)	Spalte (column)	vertikaler Spaltenindex einer Tabelle
Primärschlüssel (primary key)		eindeutiger Identifikator
Fremdschlüssel (foreign key)		Schlüssel aus einer anderen Tabelle, um eine Beziehung I

Quelle: https://de.wikipedia.org/wiki/Relation\_(Datenbank)

## Beziehungen

Weiterhin können Verknüpfungen genutzt werden, um die Beziehungen zwischen Tabellen auszudrücken. Eine Bibliothekdatenbank könnte damit etwa folgendermaßen mit drei Tabellen implementiert werden:

Tabelle Buch, die für jedes Buch eine Zeile enthält:

- Jede Zeile besteht aus den Spalten der Tabelle (Attributen): Buch-ID, Autor, Verlag, Verlagsjahr, Titel, Datum der Aufnahme.
- Als Schlüssel dient die Buch-ID, da sie jedes Buch eindeutig identifiziert.

Tabelle Nutzer, die die Daten von allen registrierten Bibliotheksnutzern enthält:

• Die Attribute wären zum Beispiel: Nutzer-ID, Vorname, Nachname.

#### Beispiel einer Relation "Buch":

Buch-ID	Autor	Verlag	Verlagsjahr	Titel	Datum
1	Hans Vielschreiber	Musterverlag	2007	Wir lernen SQL	13.01.2007
2	J. Gutenberg	Gutenberg und Co.	1452	Drucken leicht gemacht	01.01.1452
3	G. I. Caesar	Handschriftverlag	-44	Mein Leben mit Asterix	16.0344
5	Galileo Galilei	Inquisition International	1640	Eppur si muove	1641
6	Charles Darwin	Vatikan-Verlag	1860	Adam und Eva	1862

		Relat "Entlie		
Relation "Nutzer"			Nutzer-ID	Buch-ID
Nutzer-ID	Vorname	Nachname	10	1
10	Hans	Vielschreiber	10	2
11	Jens	Mittelleser	10	3
12	Erich	Wenigleser	10	3
12	EIICII	Weiliglesei	12	5
			12	6

Außerdem braucht man eine dritte *Tabelle Entliehen*, die Informationen über die Verfügbarkeit des Buches enthält.

- Jede Zeile dieser Tabelle Entliehen ordnet einer Nutzer-ID eine Buch-ID zu.
- Als Schlüssel nimmt man hier die Attributmenge (Nutzer-ID, Buch-ID).

Gleichzeitig verbindet die Nutzer-ID jeden Eintrag der Tabelle Entliehen mit einem Eintrag der Tabelle Nutzer, sowie die Buch-ID jeden Eintrag von Entliehen mit einem Eintrag der Tabelle Buch verbindet. Deswegen heißen diese Attribute in diesem Zusammenhang Fremdschlüssel (engl. foreign key).

Der hier benutzte Begriff Relation beschreibt nicht die Beziehung zwischen Entitäten (wie im Entity-Relationship-Modell), sondern die Beziehung der Attribute zum Relationennamen. **Schlüssel (Key)** 

- Primärschlüssel (primary key)
  - ein Attribut oder Attributkombination einer Relation, die einen Datensatz einer Tabelle eindeutig identifizieren
- Fremdschlüssel (foreign key)
  - ein Attribut oder eine Attributkombination einer Relation, welches auf einen Primärschlüssel einer anderen Relation verweist (Wikipedia)

Quelle: https://de.wikipedia.org/wiki/Relationale\_Datenbank

## Gegenüberstellung von Grundbegriffen (Relationale Datenbank und ERM)

Relationale Datenbank	ERM
Kopfzeile	Entitätstyp
Spalte	Attribut
Tabelle	Entitätsmenge
Spaltenüberschrift	funktionale Beziehung (Relationship)
Zeile	Entität

Quelle: https://de.wikipedia.org/wiki/Relationale\_Datenbank

#### **Datentypen**

Alle Attributwerte eines Attributes sind von einem bestimmten Datentyp.

Attributwerte sollten nach Möglichkeit elementar (nicht weiter zerlegbar) sein.

Name	Beschreibung
integer	Ganzzahlen
double precision (float8)	Gleitkommazahl
character varying (varchar)	Zeichenkette (variable Länge)
text	Zeichenkette
date	Datum
timestamp	Datum und Zeit
boolean	Boolsche Variable (TRUE/FALSE)

## Datenbanksprache (SQL)

SQL-Introduktion: https://sqlbolt.com/

## Geodatenbanken

Geodatenbanken unterstützen in logischer und interner Schicht räumliche...

- Objekte
- Abfragen
- Bezugssystemen

Konventionelle Datenbanken können Geodaten nicht effizient verwalten. Daher gibt es für viele Datenbanken Erweiterungen für die Verwaltung von Geodaten. (Wikipedia)

#### **Problem**

- 1. Geometrien mithilfe von atomaren Datentypen in einer DB abzuspeichern ist aufgrund der Eigenschaften von Geometrien nicht effizient
- 2. Datenbank interne Suchmechanismen mithilfe von Indexstrukturen sind meist auf eindimensionale Daten ausgelegt

#### Lösung

- 1. Geodatenbanken (z.Bsp. PostGis) besitzen spezielle Datentypen für Geometrien
- 2. Geodatenbanken nutzen mehrdimensionale (räumliche) Indexstrukturen
  - Oder überführen räumliche Objekte in einen eindimensionalen Raum, sodass herkömmliche Indexstrukturen verwendet werden können

#### Indexstrukturen

- Über grundlegende Dateiorganisationsform hinausgehende Zugriffsstruktur zur Effizienzverbesserung
- Zugriffspfad: Datenstruktur für zusätzliche, schlüsselbasierten Zugriff auf Tupel
- eindimensionale Indexstrukturen
  - B+ Baum
- mehrdimensionale (räumliche) Indexstrukturen
  - Quadtree, R-Tree, GIST

## **OpenGIS - Simple Feature Specification (OGC)**

Simple Feature Access ist eine Spezifikation des Open Geospatial Consortium, welche eine allgemein gültige Architektur für geografische Daten und deren Geometrien definiert.

Die Spezifikation definiert für DBMS:

- · räumliche Datentypen
- räumliche Datenstrukturen
- räumliche Funktionalitäten

#### Geometrie Klassenmodell

Folgende instanziierbare Klassen beinhaltet das Modell (Wikipedia)

- Punkte (Point)
- Linien (LineString)
- Polygone (Polygon)
  - Hierbei sind die Punkte des äußeren Ringes entgegengesetzt dem Uhrzeigersinn (linksdrehend) sortiert.
  - Die Punkte des inneren Ringes (Loch) im Uhrzeigersinn (rechtsdrehend) sortiert
- Mehrere Punkte (MultiPoint)
- Mehrere Linien (MultiLineString)
- Mehrere Polygone (MultiPolygon)
- Sammlung dieser Geometrien (GeometryCollection)

### Repräsentation der Geometrie (Well-known Text (WKT))

Die Well-known-Text-Repräsentation wird vor allem dazu verwendet, um die Geometrie alphanumerisch darstellen zu können (Wikipedia)

#### **Punkt**

```
Point(10 10)
```

**LineString** (Linie mit "Knickpunkten")

```
LineString (10 10, 20 20, 30 40)
```

#### Polygon (Fläche)

ein Polygon wird von zwei Klammern umschlossen. Polygon ohne Loch:

```
Polygon((10 10, 10 20, 20 20, 20 15, 10 10))
```

mit einem äußeren Ring und einem inneren Ring (Loch)

```
Polygon((0 0, 0 20, 20 20, 20 0, 0 0),(5 5, 5 15, 15 15, 15 5, 5 5))
```

#### Mehrfachpolygon

zwei Polygone

```
MultiPolygon(((10 10, 10 20, 20 20, 20 15, 10 10)),((60 60, 70 70, 80 60, 60 60 )))
```

zwei Polygone, erstes Polygon mit Loch:

```
MultiPolygon(((0 0, 0 20, 20 20, 20 0, 0 0),(5 5, 5 15, 15 15, 15 5, 5 5)),((30 30, 30 40, 40 40, 40 30, 30 30)))
```

#### räumliche Funktionen

Siehe auch "PostGIS Special Functions" in der PostGIS Dokumentation: https://postgis.net/docs/PostGIS\_Special\_Functions\_Index.html

#### **PostGIS Fuctions**

- ST\_Area
  - Returns the area of the surface if it is a Polygon or MultiPolygon.
- ST\_Buffer
  - (T)Returns a geometry covering all points within a given distancefrom the input geometry.
- ST\_Centroid
  - Returns the geometric center of a geometry.
- ST SRID
  - Returns the spatial reference identifier for the ST\_Geometry as defined in spatial\_ref\_sys table.

• ...

## Verschneidung

- Intersection():Geometry
- Union():Geometry
- Difference():Geometry
- SymDifference():Geometry

#### Test

- Equals():Integer
  - Gleichheit mit einer anderen Geometrie
- Disjoint():Integer
  - Menge der gemeinsamen Punkte = 0
- Intersects():Integer
  - Schnitt mit einer anderen G.
- Touches():Integer
  - Berührung mit einer anderen G.
- · Crosses():Integer
  - Kreuzung mit einer anderen G.
- Within():Integer
  - innerhalb einer anderen Geometrie
- · Contains():Integer
  - Enthält eine andere G.
- Overlaps():Integer
  - Überlappung mit einer anderen G.
- · Relate():Integer
  - Beziehung mit einer anderen G.

# **Geodateninfrastruktur (GDI)**

### **Begriffe**

## Komponenten

- Nationale Geodatenbasis (NGDB)
  - Geobasisdaten (GBD)
    - \* grundlegende amtliche Geodaten, welche die Landschaft (Topographie), die Flurstücke und die Gebäude anwendungsneutral beschreiben
  - Geofachdaten (GFD)
    - \* raumbezogene Daten aus einem Fachgebiet

- Metadaten (MD)
  - \* Informationen über andere Daten
- GI-Netzwerke
- GI-Dienste
  - Funktionen, die aus Geodaten kontextabhängige Geoinformationen produzieren
- Standards

#### Merkmale

- 1. Dezentralisierung von Geodaten und GI-Diensten
- 2. Technische-organisatorische Infrasturktur
- 3. Internationale Standards für Schnittstellen (Interoperabilität)
- 4. Konsensgetriebene Organisationstruktur

#### **INSPIRE**

## INfrastrucuter for SPatial InfoRmation in Europe

EU-Beschluss einer Richtlinie zum Aufbau einer europäischen Geodateninfrastruktur innerhabl der EU

#### **Ziele**

- Interoperabilität durch Spezifikationen
  - einheitliche Geodaten und Metadaten in Europa
  - Verfügbarkeit für verschiedene Nutzergruppen
  - Nutzung internationaler Standards

#### **Standards**

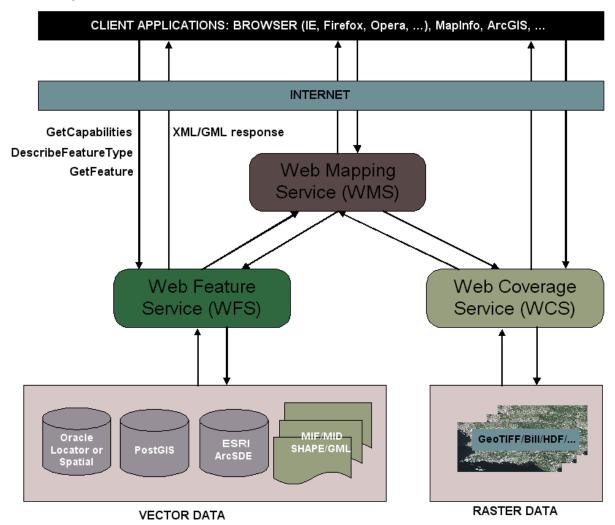
- Open Geospation Consortium (OGC)
  - internationales Konsortium aus Behörden, Industrie und Universitäten
- International Organizatin for Standardization (ISO)
  - offizielle Standardisierungsorganisation
  - in Deutschland durch DIN vertreten
  - GIS-Bereich: Arbeitsgruppe (TC) 211
- · CEN/ DIN
  - im GIS-Bereich sind Normen durch ISO-Normen abgelöst

## **Open Geospatial Consortium (OGC)**

Das OGC ist ein internationales Konsortium aus Behörden, Industrie und Universitäten. Das OGC hat bereits zahlreiche Standards veröffentlicht (Z.Bsp. OpenGIS Web Services (OWS)).

## **OpenGIS Web Services (OWS)**

- Web Feature Servie (WFS)
  - Schnittstelle zum Abrufen von Geodaten über das Internet
  - beschränkt sich auf Vektordaten, wie sie in der Datenbank abgelegt worden sind
- Web Map Service (WMS)
  - Schnittstelle zum Abrufen von Auszügen aus Landkarten über das Internet
  - visualisiert (Zuschneidung und inhaltliche Aufbereitung) Karten aus Rasterdaten oder Vektordaten
  - gibt meist Raster-Grafikformate zurück



# **Graphen und Netzwerkanalyse**

## **Topologie**

### **Topologie**

- Beschreibt die gegenseitige Lage und Anordnung geometrischer Gebilde im Raum
  - Beschreibung der Relationen zwischen elementaren Objekten (Knoten, Kante, Masche)
- charakterisiert nichtmetrische, räumliche Beziehungen zwischen Geoobjekten (Nachbarschaftsbeziehungen)
  - Ermöglicht die koordinatenfreie Speicherung von Objekten

### **Topologische Invarianten**

- Ein Knoten ist Endpunkt einer Kante
- Zwei Kanten kreuzen sich/ sind kreuzungsfrei
- Ein Punkt liegt im Inneren einer Fläche
- Ein Punkt liegt auf dem Rand einer Fläche
- Eine Fläche hat ein Loch
- Eine Fläche ist zusammenhängend/ nicht zusammenhängend
- · Zwei Flächen sind benachbart

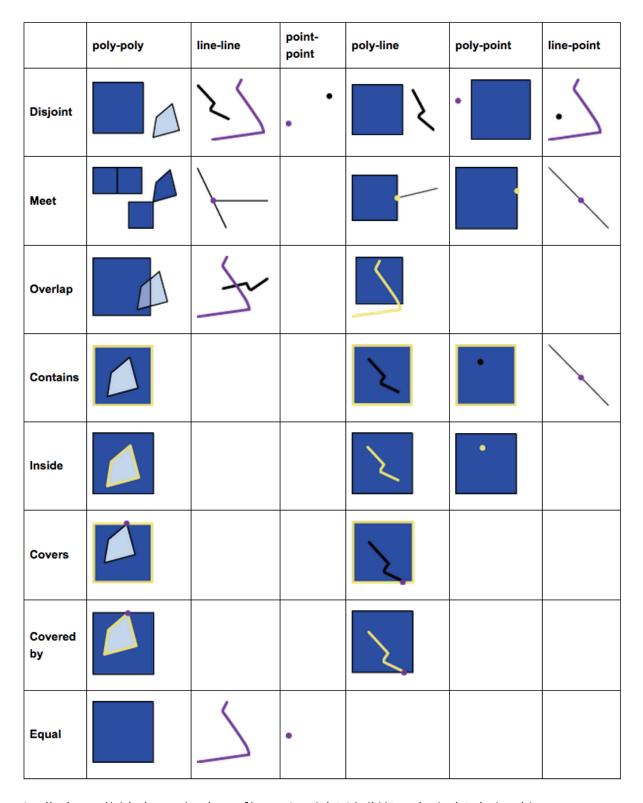
## Nicht-topologische Eigenschaften

- Abstand
- Fläche
- Winkel
- Umfang
- Durchmesser

## Topologische Beziehung nach Egenhofer

## **Egenhofer Relationen**

- · A disjoint B
- A contains B
- A inside B
- A equals B
- · A touches B
- A covers B
- · A is covered by B
- A overlaps B



Quelle: https://github.com/rugbyprof/4553-Spatial-DS/wiki/Topological-Relationships

# Graphentheorie

- Die Graphentheorie liefert ein mathematisches Fundament für die einheitliche Darstellung und Analyse unterschiedlicher Netzwerke.
- Daher beschränkt sich ein Graph vorwiegend auf die Modellierung von Beziehungen seiner Elemente, der Konnektivität.
- Konnektivität ist eine topologische Eigenschaft

## Graph

- Ein Graph ist ein Gebilde aus Knoten (nodes) und Kanten (edges)
  - Eine Kante verbindet immer 2 Knoten
  - Diese Knoten sind die Endpunkte der Kante

## **Adjazenz**

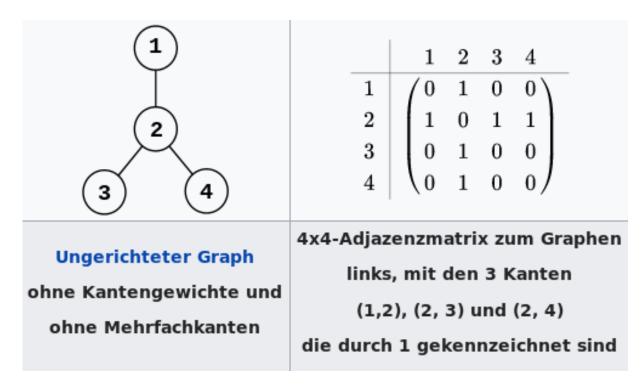
- Aneinandergrenzen gleichartiger Primitive
  - Z.B. 2 Knoten über eine Kante, 2 Kanten über einen Knoten

#### Inzidenz

- Aneinandergrenzen unterschiedlicher Elemente
  - Z.B. Knote und Kante

# Adjazenmatrix (Nachbarschaftsmatrix)

- eine Matrix, die speichert, welche Knoten des Graphen durch eine Kante verbunden sind (Wikipedia)
- Sie besitzt für jeden Knoten eine Zeile und eine Spalte, woraus sich für n Knoten eine n×n Matrix ergibt



Quelle: https://de.wikipedia.org/wiki/Adjazenzmatrix

# ungerichteter Graph

• Es werden lediglich die Verbindungen zwischen den Knoten dargestellt (Verbindungsschema)

# gerichteter Graph

• Es wird zusätzlich dargestellt, in welcher Richtung Verbindungen zwischen den Knoten bestehen

# gewichteter Graph (bewerteter)

• Graph, dessen Kanten mit Gewichten bewertet sind (z.B. den Kantenlängen)

# zusammenhängender Graph

• wenn für zwei beliebige Knoten (mindestens) ein Weg besteht

# vollständiger Graph

• wenn alle Knotenpaare adjazent sind

#### **Schleife**

• eine gerichtete Kante in einem gerichteten Graph, die einen Knoten mit sich selbst verbindet

#### **Baum**

• schleifenloser, zusammenhängender Graph, in dem je zwei beliebige verschiedene Knoten durch genau einen Weg verbunden sind

#### Netzwerkanalyse

## Grundprobleme

- 1. Bester-Weg-Problem
- 2. Bester-Standort-Problem
- 3. Problem des Handlungsreisenden (Traveling Salesman Problem)
  - Die Aufgabe besteht darin, eine Reihenfolge für den Besuch mehrerer Orte so zu wählen, dass keine Station außer der ersten mehr als einmal besucht wird, die gesamte Reisestrecke des Handlungsreisenden möglichst kurz und die erste Station gleich der letzten Station ist (Wikipedia).

#### GIS Funktionen für Netzwerke

GIS beinhalten Werkzeuge zur Erstellung korrekter Netztopologien.

- Unterstützung von Kantengewichten
- · Katenübergänge (Turns)
  - An Knoten bestehen mehrere Möglichkeiten, von einer Kante zur anderen zu wechseln
  - Modellierung von Abbiegevorschriften und Stops
- Einzugsgebiet (Service Area)
- · Erreichbarkeitsanalyse (Isochronen)
  - Berrechnung eines Polygons, das das innerhalb einer vorgegeben Zeit zu erreichende Gebiet repräsentiert

#### Lokation-Allokation-Modellierung

- Location-allocation refers to algorithms used primarily in a geographic information system to determine an optimal location for one or more facilities that will service demand from a given set of points (Wikipedia)
- Wenn Einrichtungen, die Waren und Services bereitstellen, und verschiedene Bedarfspunkte, welche diese nutzen, gegeben sind, dann besteht das Ziel der Location-Allocation darin, diejenigen Einrichtungen zu suchen, die die Bedarfspunkte am effizientesten bedienen. Wie der Name bereits sagt, bezeichnet Location-Allocation eine Problemstellung mit zwei Aspekten, wobei gleichzeitig Einrichtungen gesucht und Bedarfspunkte den Einrichtungen zugeordnet werden. (ArcMap Hilfe)
- z.Bsp: Viele Einzelhandelsgeschäfte beziehen ihre Güter von Industriebetrieben. Gleichgültig, ob Autos, Küchengeräte oder Nahrungsmittel produziert werden, können für einen Industriebetrieb große Anteile seines Budgets auf Transportkosten entfallen. Location-Allocation kann die folgende Frage beantworten: Wo sollte sich der Industriebetrieb ansiedeln, um die Gesamttransportkosten zu minimieren? (ArcMap Hilfe)

#### Distanzen in Netzwerken

#### **Geometrische Distanzen**

- Euklidische Metrik (Satz von Pythagoras)
- Cityblockmetrik (N4-Nachbarschaftstyp)
- Schachbrettmetrik (N8-Nachbarschaftstyp)

Was ist die Distanz zwischen 2 Polygonen?

- Minimal
- Zentroid (Schwerpunkt)

# Distanzen in Netzwerken (topolog. Graphen)

- Anzahl der Knoten eines Weges in einem Graph
- Summe der Gewichte für Wegkosten auf Kanten zwischen Start- u. Zielknoten

## Routenplanung

#### Motivation

• ist der Wunsch, optimal im Sinne seiner persönlichen Bedürfnisse und der gewählten Fortbewegungsart von seinem Standort zu einem bestimmten Zielort zu gelangen.

#### **Bedüfnisse**

- anzukommen
- schnell
- · wenig Energie
- wenig Gefahrenstellen
- Zwischenziele

## Fortbewegungsarten (Verkehrsmodalitäten)

- Fahrrad
- Auto
- zu Fuß
- Schiff
- Eisenbahn
- Flugzeug
- ...

## **Individualisierte Routenplanung**

• Ermittlung der Kantenkosten mit einer anwenderspezifischen Formel

```
    Kantenkosten = x * Distanz + y * Steigung + z * Panorama +...
    Für Fahrradbote: x >> y >> z
    Für Tourist: z >> y >= x
```

Siehe auch: https://maps.openrouteservice.org/

# **3D**

2D

 XY
 2.5D
 XY(Z)
 z = f(x,y)

 3D

 XYZ

# Lernziele

- Wofür ist die dritte Dimension wichtig
- Von Messungen zu 3D Objekten
- Grundlegende Datenmodelle für 3D-Geoobjekte
- Reliefanalysen
- Ableitung von 3D Stadtmodellen

# Kontrollfragen

- Welche Objektmodelle gibt es für digitale Geländemodelle?
- Wie kann man 3D Objekte repräsentieren?
- Beispiele für Reliefanalysen und Anwendungsfälle?
- Wie wird ein TIN erzeugt?
- Was sind Anwendungsfälle für 3D Stadtmodelle?
- Welche Schritte sind notwendig um aus einem Laserscan ein 3D Stadtmodel zu erzeugen?

# OpenStreetMap

"[...] OpenStreetMap, the project that creates and distributes free geographic data for the world. We started it because most maps you think of as free actually have legal or technical restrictions on their use, holding back people from using them in creative, productive, or unexpected ways." (OpenStreetMap Wiki)

#### **Datenstruktur**

#### Grundelemente

- Punkt oder Knoten (Node)
- Linie oder Weg (Way)
- Attribut (Tag)

#### **Elemente**

- Nodes
  - Punkt oder Knoten
  - Geometrischer Punkt (durch Koordinaten definiert)
- Ways
  - Aufeinander folgende Geradenabschnitte (segment)
  - Gesamte Linie (way) wird mit einem Attribut (tag) versehen
  - Haben eine Richtung
  - Können geschlossen sein
- Area
  - Eigene geschlossene Linie mit besonderen Atributen
  - (area = yes (explitzit))
  - (building = yes (implizit))
- Relations
  - Beschreibt Beziehungen zwischen Objekten oder ihren Teilen
  - Hat Attribut type = \*
  - Attribute (Eigenschaften) eines Objekts
  - Bestehen aus key und value
  - Ein Objekt kann verschiedene tags haben
  - Beispiel: eine Straße
    - \* highway = residential
    - \* name = "Hauptstraße"
    - \* maxspeed = 50

## **Datenmodell**

Problem	Vorteil
Freies Tagging Schema	Format auch für Laien leicht zu verstehen
Kein Standard	Wichtiges Argument in einem System, dass allen Menschen offen
Datenmodell zu einfacht (Probleme zu erwarten)	

## weiterführende Materialien

# OpenStreetMap Wiki

http://wiki.openstreetmap.org/

# **European Handbook of Crowdsourced Geographic Information**

https://www.ubiquitypress.com/site/books/10.5334/bax/

# Klausur - Antworten

Siehe auch: Klausur - Fragen

# WS 12/13

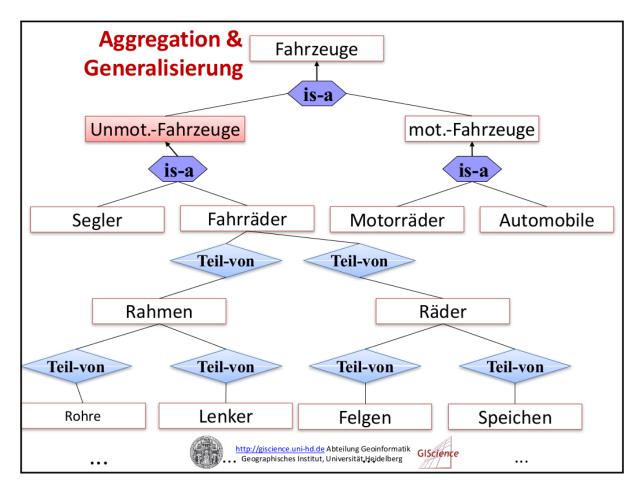
1.

- 1. Semantische Metadaten
  - Beschreibung der verwendeten Symbolik?
- 2. Syntaktische
  - Zugriffsmechanismen?

2.

Aggregation und Generalisierung bilden Hierarchien aber:

- 1. Aggregation bezieht Objekte aufeinander
- 2. Generalisierung bezieht Klassen aufeinander



#### 3.

- Entwicklung ("Lineage"):
  - die historische Entwicklung eines Datensatzes.
- Vollständigkeit ("Completeness"):
  - die Vollständigkeit eines Datensatzes.
- Logische Konsistenz ("Logical Consistency"):
  - beinhaltet sowohl topologische Konsistenz als auch die Attributierung und die Beziehungen innerhalb der DB.
- Positionsgenauigkeit ("Positional Accuracy"):
  - auf Koordinaten bezogen.
- Zeitliche Genauigkeit ("Temporal Accuracy"):
  - "Accuracy of Measurement": Genauigkeit von Zeitangaben
  - "Temporal Consistency": Korrekte zeitliche Reihenfolge der Daten.
  - "Temporal Validity": Gültigkeit der Daten an bestimmtem Zeitpunkt.
- Thematische Genauigkeit ("Thematic Accuracy"):
  - die richtige syntaktische Attributierung ("Syntactic Accuracy"), als auch die http:

//giscience.uni-hd.de Abteilung Geoinformatik Institut, korrekte Zuordnung von Geographisches Objekten zu Universität ihren Heidelberg Objektklassen ("Semantic Accuracy").

#### 4.

- DHM = Digitales Höhenmodell
  - Bei einem DHM muss immer angegeben werden, um welche Oberfläche es sich handelt,
  - z.B. DHM der Vegetationsoberfläche. DHM der Erdoberfläche, DHM der Grundwasserfläche
- DOM = digitale Oberflächenmodell (engl. digital surface model DSM)
  - repräsentiert die Erdoberfläche (Grenzschicht Pedosphäre Atmosphäre) samt aller darauf befindlicher Objekte (Bebauung, Straßen, Vegetation, Gewässer usw).
- DLM = Digitales Landschaftsmodell
  - beschreiben die Erscheinungsformen und Sachverhalte der Erdoberfläche durch geotopographische Objekte in unterschiedlichen Detaillierungsgraden.
- DGM = Digitale Geländemodelle
  - beschreibt die Geländeformen der Erdoberfläche durch eine in Lage und Höhe georeferenzierte Punktmenge
- DKM = Digitales Kartographisches (Landschafts)Modell
  - Teil von ATKIS. Diente dem Ziel verscheidener Maßstäbe automatisch aus dem digitalen Landschaftsmodell abzuleiten.

# 5.

#### **Vektor:**

Vorteil	Nachteile
Scharfe Repräsentation der Geometrie	Komplexität: komplexe Datenstruktur und resultierend d
Beliebige unregelmäßige Formen und Rauminhalte	Schwierig zu bearbeiten
Kompaktheit: Geringerer Speicherplatzbedarf	Weniger geeignet für kontinuierliche Phänomene
Maßstabsunabhängig (keine Einschränkungen bei Zoom)	
Geeignet für diskrete Phänomene	

The final advantage of vector data is that topology is inherent in the vector model. This topological information results in simplified spatial analysis (e.g., error detection, network analysis, proximity analysis, and spatial transformation) when using a vector model (geographic information system basics).

## Raster:

Vorteil	Nachteile
Technologisch schnelle und günstige Erfassung von Daten	große Dateien: großer Speicherplatz
Einfache Datenstruktur und resultierend daraus einfache Analysealgorithmen	Maßstabsabhängig (Einschränkunge
Geeignet bei kontinuierlicher Phänomenen	Probleme in der Beschreibung linier
	Weniger geeignet für diskrete Phänd
	Es gibt keine Information über die B

## 6.

- Vorteile
  - Ladensweitedatensätze vorhanden
  - aktives System (Laser)?
  - Bodenpunkte möglich (durch Lücken in der Vegetation)
  - Anschauliche Visualisierungen in Verbindung mit Luftbildern
- Nachteile
  - Attributlose Punktwolke
  - Nachträgliche Objektbildung erforderlich

## 7.

# Standorteigenschaften:

- Entfernung zu anderen Krankenhäuser
- Bevölkerungsdichte
- Entfernung zu Hauptstraßen

# 8.

- Gewichtete Graphen
- \_
- 1. Grünster Weg
- 2. Interessantester Weg (Sight seeing)
- 3. Schnellster Weg
- 4. Kürzester Weg
- 5. Der Weg mit dem grinsten Verbrauch
- 6. Der Weg mit dem wenigsten Lärm

#### WS 12/13 (zweiter Klausurtermin)

1.

siehe Klausur WS 12/13 Aufgabe 3

2.

# Klausur - Fragen

siehe auch: Klausur - Antworten

## WS 12/13

- 1. Metadaten werden in syntaktische, semantische und pragmatische Metadaten unterteilt. Nennen sie je zwei explizite geographische Beispiele pro Kategorie.
- 2. Erklären Sie die Begriffe Generalisierung und Aggregation anhand eines Beispiels.
- 3. Was sind die Qualitätseigenschaften von Geodaten (Van Oort)?
- 4. Erläutern Sie den Unterschied zwischen DGM, DHM, DOM, DLM, und DKM. Schreiben Sie die Abkürzungen aus.
- 5. \_
- Nennen Sie 4 Vorteile und 4 Nachteile des Vektor- bzw. des Rasterdatenmodells.
- Welche Vor- und Nachteile ergeben sich durch Laserscanning erfasste Daten im Rasterbzw. Vektordatenmodell.
- 6. Nennen Sie die Vor- und Nachteile des Laserscannings.
- 7. Sie sind Projektleiter einer Firma, die mit der Standortsuche für ein Krankenhaus beauftragt wird. Skizzieren Sie (in Stichworten) kurz die wichtigsten Prozessierungsschritte in ihrem GIS Projekt. (Vereinfachte Darstellung)
- 8. Einige Parameter können in die individualisierte Routenplanung miteinfließen
  - Wie kann dies am einfachsten geschehen?
  - Nennen Sie 6 derartige Attribute oder Maße.

## WS 12/13 (zweiter Klausurtermin)

- 1. Nenne mind. Sechs Qualitätsmerkmale von Metadaten
- 2. Vergleiche das Spaghetti- und Topologische Datenmodell miteinander
- 3. Kardinalität von Beziehungen mit jeweils einem bsp aus der Geographie
- 4. Erkläre Transaktion an eine Beispiel

- 5. ATKIS und ALKIS im Vergleich (Signaturenkatalog)
- 6. Nenne 3d Analysefunktionen
- 7. Standortanalyse für eine Mülldeponie

#### **SS 16**

- 1. Geben Sie eine kurze Beschreibung zu den Begriffen thematische, geometrische und topologische Abfragen und formulieren Sie jeweils ein Beispiel.
- 2. Nennen Sie die Operatoren der geometrischen Verschneidung.
- 3. Welche zwei Interpretationsmöglichkeiten gibt es für den englischen Begriff "Overlay" in GIS/-Kartographie.
- 4. Was ist der Unterschied zwischen Geocoding und Reverse Geocoding?
- 5. Wie lauten die Operationen bei Tomlins Map Algebra?
- 6. Was ist der Unterschied von objektrelationalen Datenbanken im Vergleich zu relationalen Datenbanken? Was bedeutet das für Geodaten?
- 7. Was bedeuten die Begriffe Aggregation und Generalisierung (oder was ist der Unterschied zwischen Aggregation und Generalisierung?) und nennen Sie je ein Beispiel.
- 8. Erklären Sie das Prinzip (?) der "individuellen Routenplanung".
- 9. Wie hängen Delaunay-Triangulation und Voronoi-Diagramme miteinander zusammen?
- 10. Nennen Sie die Qualitätsparameter von Geodaten nach der ISO-Norm 19113. Geben Sie jeweils ein Beispiel dazu an.
- 11. Was drückt Tobler's First Law of Geography aus? (Aufgabe aus Gastsitzung)
- 12. Nennen Sie die Teilgebiete der Räumlichen Analyse. (Aufgabe aus Gastsitzung)
- 13. Nennen Sie Vor- und Nachteile des Spaghetti-Modells für Vektordaten.

## Klausur - Themen

#### **Analyse Vektor**

- Puffer
- · Thiessen/Voronoi
- Delauny
- Entfernungsmatrix/ Entf.berechnung
- Räumliche Aggregation
- · Geoprozessierung vs. Analyse
- · Rubber-Sheeting

## **Analyse Raster**

- Grundlagen Raster/Fernerkundungsdaten: Typen, Aufbau, Werte,
- Map Algebra, Grundlagen, Längen, Linien/Flächen, Typen der Operatoren
- Nachbarschaften (Bsp. Analyseablauf (Prinzip))

## Räumliche Analysen

- Grundlagen wie in Vorlesung (Grundidee erklären)
  - z.B. Toblers 1. Gesetz / 1.st Law of Geography, Interpolationsverfahren und Tuningparameter (keine Formeln)
- Unterschied Interpolation Kerndichteberechnung
- Autokorrelation; Maße für räuml. Autokorrelation, lokale und globale Autokorrelation
- Wie lassen sich Punktmuster beschreiben (Grundidee)
- Beispiele aus Geographie
- Keine Formeln

# **Netzwerke, Graphen, Routing**

- Typen von Netzwerkanalysen (bester Weg, beste Standorte (unter Nutzung von Einzugsgebieten), Idee TSP),
- GIS-Funktionen f. Netzwerke
- Distanzfunktionen
- · Adjazenzmatrix,
- Kantengewichte
- Idee der individualisierten Routenplanung

NICHT: Qualität von OSM, OSM Anwendungen, Bsp. (OSM) Routing Anwendungen, Exkurse

# Topolog. Datenstrukturen (Topol./Spagh.)

- Topologische Beziehungen nach Egenhofer (Prinzip der Formalisierung, nicht Tabelle auswendig)
- Prinzip Spaghetti vs. Topolog. Vektordatenstrukturen, Vor-Nachteile, Fehler
- Einfachste Variante Spaghetti-Modell, einfache Verbesserungsmöglk.
- Idee d. Topolog. Datenstrukturen (welche Informationen müssen wie gespeichert werden)
  - Knoten, Kanten aus Knoten mit Topologie zu Flächen rechts/links, i.d.R. extra Flächentabelle mit Liste der Kanten, Koordinaten nur bei Punkten/Knoten Rest Exkurs

#### Geodatenbanken

- Grundlegende Definitionen
- Motivation DB
- Funktionalität DB
- GeoDB
- · Eigenschaften DB
- · Transaktion, ACID
- Relationale DB, Prinzipien (nicht Historie oder Software Marken)
- Suchindex in GeoDB (warum?), Prinzip Quadtree/R-Tree
- Objektrelationale DB (Prinzip)
- Wesentliche Geometrietypen nach Simple Feature Specification (OGC SFS)
- GIS-Analyseoperationen in OGC SFS möglich (welche Type? Keine Syntax)

Nicht: SQL

# **Datenmodellierung**

- Datenmodellierung, Grundlagen, Phasen
- Grundlagen ER
- Modell, Elemente, Vor-Nachteile, Kardinalitäten,
- Grundlagen bis Aggregation/Generalisierung (Achtung: hier nicht "kartographische Generalisierung", sondern bei Datenmodellierung)

Nicht: Normalformen, nicht UML

#### 3D-GIS

- DGM, DHM, DOM...
- TIN
- Delauny, Voronoi/Thiessen-Analyse
- und Nutzungsmöglk. von DGM-Stadtmodelle Nutzungsmöglkeiten,
- Verfahren zur Erfassung von (3D) Geodaten
- Räuml. Repräsentationen (3D)
- · Typ. Repräsentationen für DGM

NICHT: Spezifika 3D Analyst/ArcScene

ABER: Slope, Aspekt, Contours, Hillshading, Sichtlinien/ Sichtfelder/ Steilster Pfad/ Profile etc. sind allgemeine DGM-Operationen!

# GeoWeb / GDI:

- GDI Formel erläutern können: GDI := Geodatenbasis (NGDB=GBD+GFD+MD) + GI ☑ Netzwerk + GI-Dienste + Standards
- · Geobasisdaten vs. Fachdaten
- Was ist Inspire
- Wichtigste OGC Web Services: WMS, WMTS, WFS (je min1 Satz Def.)
- Was macht WMS (NICHT AnfrageDetails, Keine Syntax)

# **Dasymetric Mapping**

- MAUP
- Areal Interpolation (Prinzip, keine Formel)
- Prinzip Dasymetric Mapping

NICHT: GIS Health and Disaster Management

## Laserscanning

- Messung der Geländehöhe durch Laufzeitmessung der Laserpulse?
- Ergebnis sind Oberflächenmodelle (2,5D)
- Attributlose Punktwolke
- Nachträgliche Objektbildung erforderlich
- 1. Grundlagen
- 2. Attributdaten (Sachdaten) und Geodaten
- 3. Datenmodelle
  - Rastermodelle
  - Vektormodelle
- 4. Datenbanken
- 5. Geospatial Analysis
  - Vektor Operationen
  - · Raster Operationen
- ?. Abfrage

# Räumliche Analyse - Interpolation

## **Begriffe**

## Interpolation

- Eine Klasse von mathematischen Problemen und Verfahren
- Zu gegebenen diskreten Daten (z. B. Messwerten) soll eine stetige Funktion (die sogenannte Interpolante oder Interpolierende) gefunden werden, die diese Daten abbildet (Wikipedia

## **Räumliche Interpolation**

- Vom Punkt (oder Linie) in die Fläche
- Eine flächenhafte Aussage aus eine Stichprobe generieren
- Die Oberflächeninterpolationswerkzeuge bilden aus Referenzpunktwerten eine kontinuierliche (oder vorhergesagte) Oberfläche ArcGIS Tools

#### Lineare räumliche Interpolation

- Linearkombination der Werte der Beobachtungen
- Summe der Gewichte ergibt 1

#### **Toblers 1. Gesetz**

• "Everything is related to everything else, but near things are more related than distant things"

#### **Nachbarschaften**

- k-nächste Nachbarn
- distanzbasiert
- k-nächste Nachbarn, Je Quadrant
- · weitere Konzepte
  - Topologische Nachbarschaft
  - Interaktionen (Austausch von Gütern, Informationen, ...)
  - Konzeptionelle oder soziale Distanz

## Klassifikation von Interpolationsverfahren

- Global
  - Modell berücksichtigt alle Punkte
- Lokal
  - nur Werte in der "Nachbarschaft" werden berücksichtigt
- Exakt
  - Beobachte Werte werden exakt getroffen (falls Feature oder feines Raster)
- Nicht-Exakt
  - Vorhergesagte Werte stimmen nicht mit den beobachteten überein

- Deterministisch
  - keine Information über den Fehler
- Stochastisch
  - probabilistische Schätzgröße

#### Interpolationsverfahren

#### Thiessen Polygone (Voronoi Diagramm)

- Unterteilung eines Gebietes, für das nur punkthafte Informationen vorliegen, in Polygone nach dem Kriterium der kürzesten Distanz zum nächsten Punkt (Spektrum)
- · Jeder Punkt erhält den Wert seines nächsten Nachbarn zugeordnet
- Harte Übergänge

#### Animation (Voronoi Diagram and Delaunay Triangulation) from miyu

• https://github.com/miyu/voronoi/blob/master/images/result.gif?raw=true

#### **Weitere Informationen**

- https://pro.arcgis.com/de/pro-app/tool-reference/analysis/create-thiessen-polygons.htm
- http://www.gitta.info/Accessibilit/de/html/UncProxAnaly\_learningObject4.html

# **Dreicksvermaschung (Triangulated Irregular Network (TIN)**

Dreiecksvermaschungen (TIN – Triangulated Irregular Network) verbinden die gemessenen Geländepunkte mit einem Netz von Dreieckskanten. Die Geländeoberfläche wird durch ein Dreieckspolyeder approximiert. Gebräuchlich für topographische Anwendungen ist die modifizierte Methode nach Delaunay, die als Zwangsseiten Geländelinien berücksichtigen kann. Dies ist aus morphologischen Gründen notwendig. Dreiecksvermaschungen werden z. B. verwendet, wenn gemessene DGM eine inhomogene Punktverteilung aufweisen. Dies ist z. B. bei digitalisierten Höhenlinien oder Echolotprofilen der Fall. Durch die lineare Verbindung der gemessenen Punkte können in diesen Fällen Artefakte weitgehend vermieden werden (Spektrum)

#### **Weitere Informationen**

- https://pro.arcgis.com/de/pro-app/help/data/tin/tin-in-arcgis-pro.htm
- https://www.e-education.psu.edu/geog486/node/1875

## **Inverse Distance Weighted (IDW)**

- Alle Beobachtungen in der Nachbarschaft gehen in die Berechnung ein
- Gewichtung ist abhängig von der inversen Entfernung (Summer der Gewichte = 1)
- keine Berücksichtigung von Covariaten
  - müssen separat berücksichtigt werden
  - gilt für die meisten Interpolationsverfahren

## Potenzt (Power)) (ArcGIS - References)

- IDW beruht hauptsächlich auf der Inversen einer potenzierten Entfernung
- Mithilfe des Parameters Potenz können Sie die Bedeutung bekannter Punkte für interpolierte Werte auf der Basis von deren Entfernung zum Ausgabepunkt steuern
- Es ist eine positive, reelle Zahl mit dem Standardwert 2
- Durch Festlegen eines höheren Potenzwerts kann die Bedeutung der am nächsten gelegenen Punkte verstärkt werden (Je höher die Potenz, desto mehr nähern sich die interpolierten Werte an den Wert des am nächsten gelegenen Referenzpunktes an)

#### **Weiter Informationen**

- https://pro.arcgis.com/de/pro-app/tool-reference/3d-analyst/how-idw-works.htm
- https://www.e-education.psu.edu/geog486/node/1877

_			
Work in Progress (wip)	)		
Kriging			
IVI 181118			

## **Sonsiges**

**Tuning Parameter** 

Bedeutung von Trend und Covariaten für die Interpolation

# **Räumliche Interpolation**

## **Varianten lokaler Interpolation**

linear gekrümmt

## Nachbarschaft (räumlich)

- k-nächste Nachbarn
- distanzbasiert
- k-nächste Nachbarn je Quadrant

Es existieren weitere Nachbarschaftskonzepte

- Topologische Nachbarschaft
- Interaktion
- konzeptionelle oder soziale Distanz

#### Klassifikation von Interpolationsverfahren

- Global Modell berücksichtigt alle Punkte
- Lokal
  - nur Werte in der "Nachbarschaft" werden berücksichtigt
- Exakt
  - Beobachte Werte werden exakt getroffen (falls Feature oder feindes Raster)
- Nicht exakt
- Deterministisch
  - keine Information über den Fehler
- Stochastisch
  - probalistische Schätzgröße

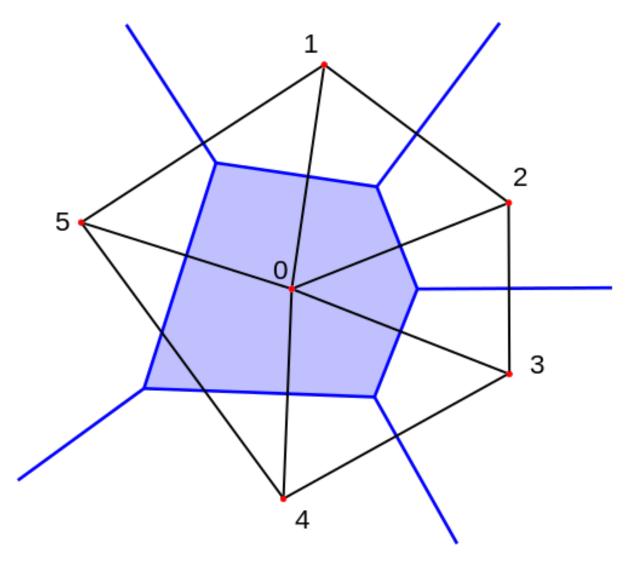
# **Bekannte Optionen**

## Dreiecksvermaschung

• ...

# Thiessen Polygone/ Voronoi Diagram

- Jeder Punkt erhält den Wert seines nächsten Nachbarn zugeordnet
- Harte Übergänge



# **Natürliche Nachbarn Interpolation**

- Thiessen Polygone für alle Beobachtungen
- Verschneiden mit Thiessenpolygonen ...

## Inverse-Distance-Weighted-Interpolation (IDW)

Inverse Distance Weighting (IDW) bezeichnet eine Interpolationsmethode, der die Annahme zugrunde liegt, dass die punktuell gemessenen raumbezogenen Daten in Abhängigkeit von der Distanz im Raum gewissen Ähnlichkeiten in den Werten aufweisen. Bei der IDW wird der Wert an einem nicht beobachteten Punkt z(u0) durch ein gewichtetes Mittel der benachbarten gemessenen Werte der Beobachtungsvariable geschätzt. Die Gewichte des dabei verwendeten linearen Schätzers sind proportional zu den Inversen des Abstands die zwischen dem unbekannten Punkt und dem verwendeten

gemessenen Punkt. (Geoinformatik - Uni Rostock)

- Alle Beobachtungen in der Nachbarschaft gehen in die Berechnung ein
- Abnahme des Gewichtes mit der Distanz

#### \$ Formel \$

- Parameter1
- Paramet2
- ...
- p
- Power
- hoher Wert -> Übergang zu Thiessen Polygone/ Voronoi Diagram
- n
- Anzahl von Nachbarn
- Mehr Nachbarn berücksichtigt -> glattere Oberfläche
- r
- Radius

# **Einbeziehung von Covarianz** (Mit Einfluss der Geländehöhe)

• ..

# **Kriging - Geostatistik**

Überbegriff für ein Set von verschiedenen Methoden.

http://wiki.gis.com/wiki/index.php/Kriging

# Ablauf (Schematisch) zur Gewichtsbestimmung

- ...
- Allgemeines
- Einführung
- Analysemethoden
- Rasterdaten
- Vektordaten
- Datenmodellierung
- Datenbanken
- Relationale Datenbanken

- Geodatenbanken
- Graphen- und Netzwerkanalysen
- Geodateninfrastruktur
- Exkurs OpenStreetMap
- Klausur
- Klausur Fragen
- Klausur Antworten