

Sichtbarmachung von Geo-Daten für eine Suchmaschine

Frauke Heinzle
Monika Sester

Institut für Kartographie und Geoinformatik (ikg)
Universität Hannover

‘Roter Faden’

- ▶ EU-Projekt SPIRIT – kurze Einleitung
- ▶ Problemstellung
- ▶ „klassische“ Metadaten
- ▶ Extraktion von impliziter Information aus Geodaten
- ▶ Beispiele und Ausblick

EU-Projekt SPIRIT – Spatially-Aware Information Retrieval on the Internet

SPIRIT – Einleitung

Problemstellung

Metadaten

Explizite Information

Beispiele, Ausblick

Ziel des Projekts:

Entwicklung einer Internet Suchmaschine, welche
räumliche Strukturen berücksichtigt

- ▶ Erweiterung der Konzepte existierender Suchmaschinen
- ▶ Betrachtung des räumlichen Kontext einer Anfrage
- ▶ Wie kann man räumliche Strukturen computergerecht beschreiben?

www.geo-spirit.org

Beschreibung räumlicher Strukturen computergerecht?

PIRIT – Einleitung

Problemstellung

Metadaten

explizite Information

Beispiele, Ausblick

- ▶ Ontologien
 - Thesauri
 - GIS-Daten (explizite und implizite Information)
- ▶ “klassische” Metadaten
- ▶ Erkennen des impliziten Wissens in Geo-Datensätzen
 - interessierende Konzepte für SPIRIT
 - Speicherung der Information

Metadaten – Daten über Daten

PIRIT – Einleitung

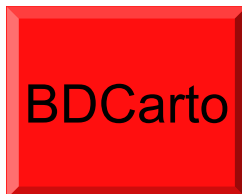
Problemstellung

Metadaten

explizite Information

Beispiele, Ausblick

- ▶ typische Standards: Dublin Core, ISO/DIS 19115
- ▶ Inhalte:
 - Name
 - Ausdehnung
 - Beschreibung etc.
- ▶ Aussagekraft
- ▶ manuelle Anreicherung
- ▶ Beispiel



| AOBJID | OBJ | TEIL | OBJART | AKT | OART_ATYP |
|---------|---------|------|--------|-----|-----------|
| B01HW10 | B01HW10 | 001 | 3102 | | 3102 |
| B01HXUA | B01HXUA | 001 | 3101 | | 3101_1307 |
| B01HXUG | B01HXUG | 001 | 3102 | | 3102 |

Ein Bild sagt mehr als tausend Worte.....

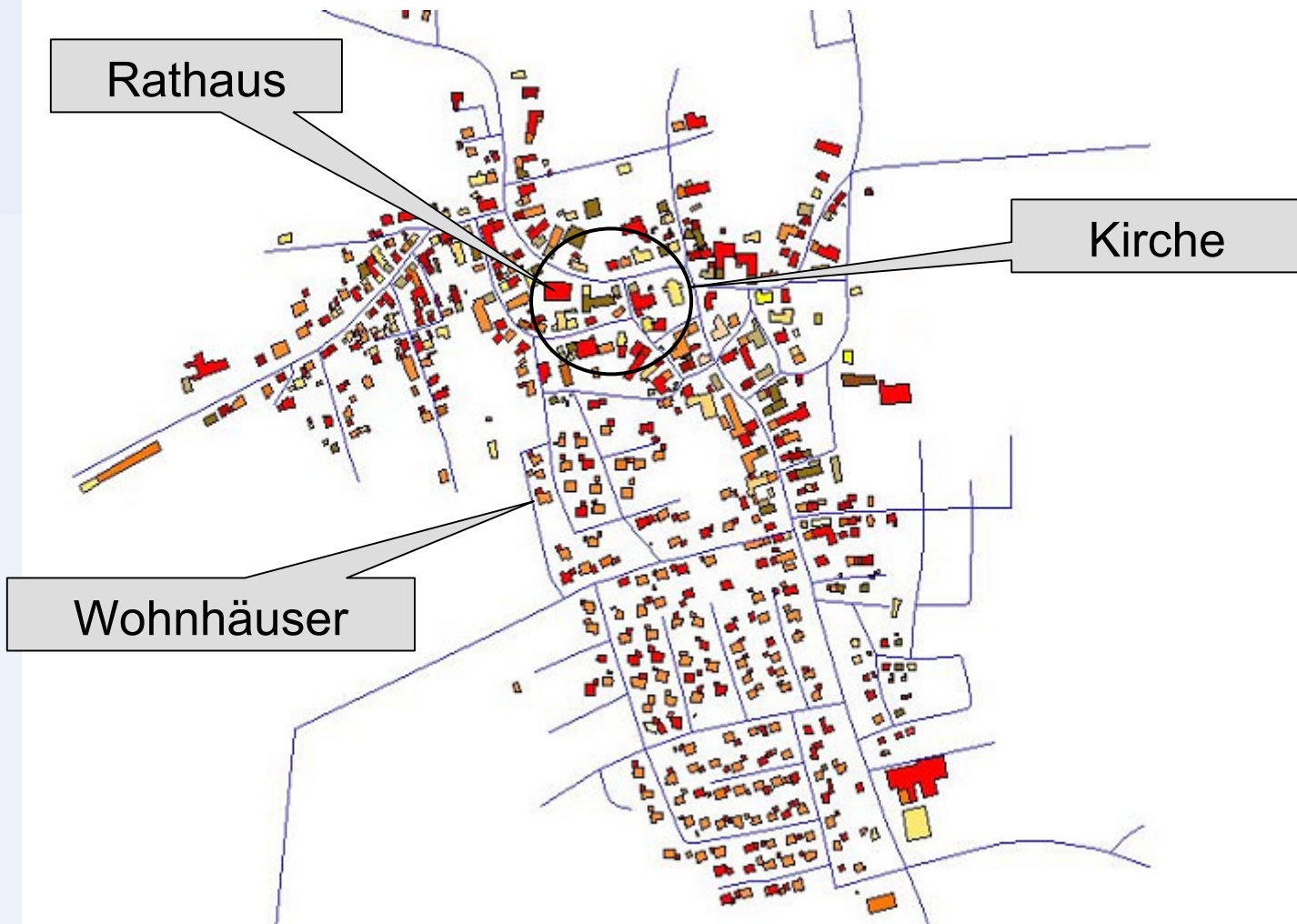
PIRIT – Einleitung

Problemstellung

Metadaten

explizite Information

Beispiele, Ausblick



Das Ganze ist mehr als die Summe seiner Teile

PIRIT – Einleitung

Problemstellung

Metadaten

Explizite Information

Beispiele, Ausblick

.....

- ▶ Inhalt ist mehr als nur pure Geometrie
- ▶ automatische Extraktion impliziter Information aus GIS-Daten
- ▶ Auffinden der geographischen Semantik
- ▶ Ableiten von räumlichen Verbindungen zwischen einzelnen Objekten
 - Topologie
 - Entfernungen
 - Richtungen
 - Häufungen von Objekten
 - typische Strukturen zwischen Objekten

- ▶ Ausdehnung eines Gebietes



- ▶ Dichte, Häufung

- ▶ Klassifizierung von verschiedenen Gebietstypen in Kategorien

- ▶ Kreuzungen von Verkehrswegen, Überlagerung von Gebieten

- ▶ Distanzauswertung

- ▶ Vorkommen und Einmaligkeit von geographischen Erscheinungen

- ▶ Methoden des clusterings

- ▶ Nachbarschaften, Aneinandergrenzen, Verbindungen bewerten

- ▶ Gebietsabgrenzung

- ▶ verschiedene Typen derselben Kategorie (Wohnbebauung, Industriegebäude, Verkehrsstrassen etc.)

- ▶ Streusiedlungen

- ▶ nahtloser Übergang zweier Städte

- ▶ Ansammlung von Gebäuden, Straßen, speziell ausgewiesenen Zonen in einer Region

- ▶ absolute, relative Distanzen

- ▶ Level of Detail - Problem

- ▶ Bestimmung von Referenzpunkten, Lokalisierung

räumliche Konzepte für SPIRIT

SPIRIT – Einleitung

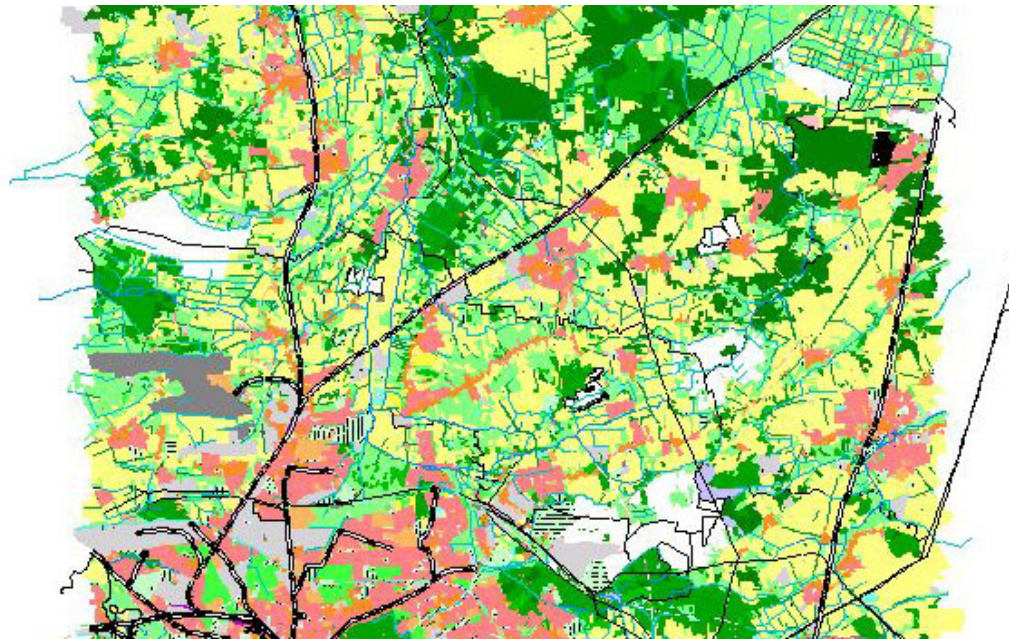
Problemstellung

Metadaten

explizite Information

Beispiele, Ausblick

- ▶ Klassifizierung von Städten (bedeutend – unbedeutend)
- ▶ Einzugsbereich von Städten
- ▶ Festlegung des Stadtzentrums
- ▶ Bestimmung von touristischen Gebieten und attraktiven Orten
- ▶ Möglichkeiten der Besiedlung oder der Ansiedlung von Industrie (Stadtentwicklung, Lebensqualität)

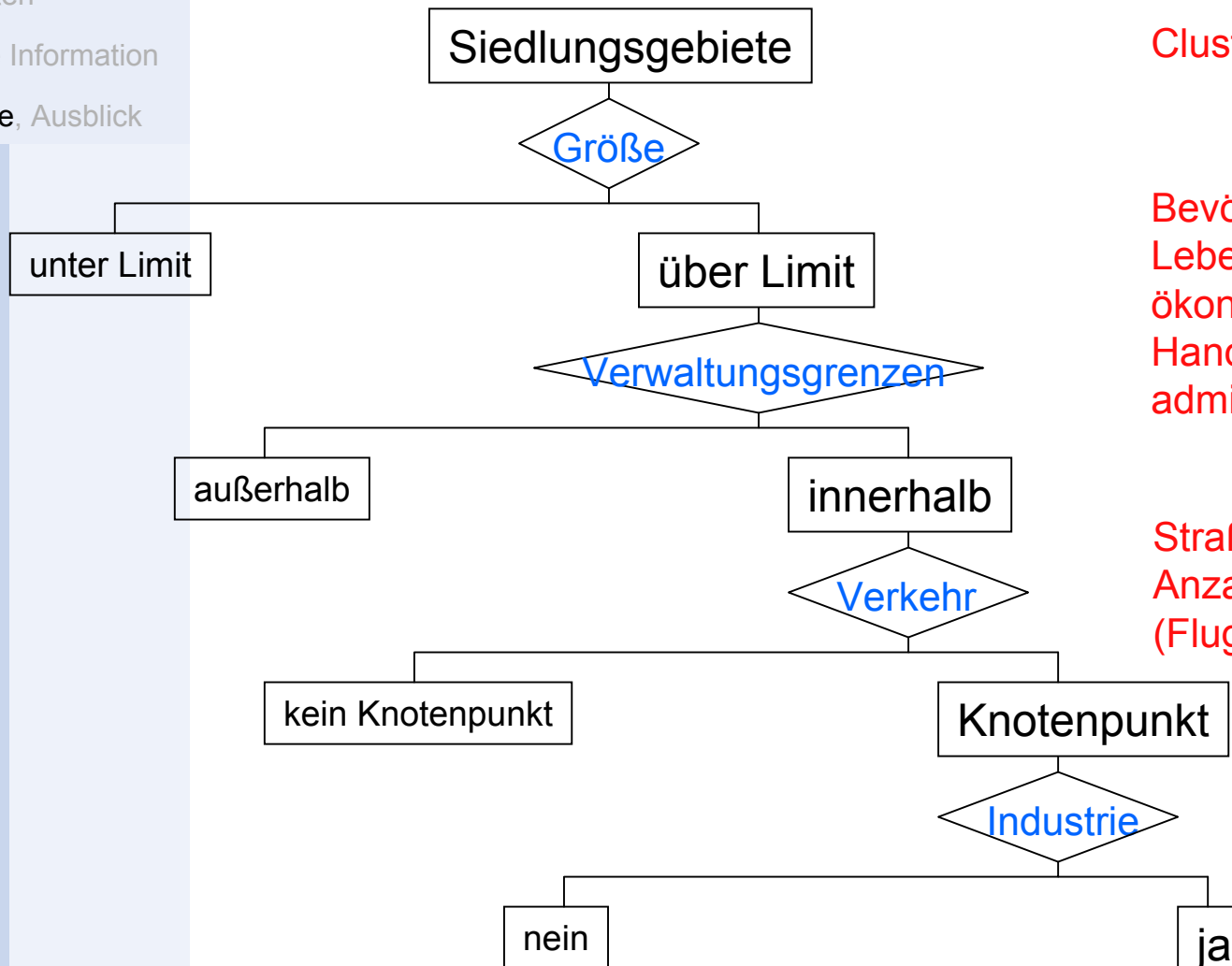


- ▶ Interpretation der GIS-Daten:
 - überwachtes Vorgehen: Modelle beschreiben die gesuchte Information (Muster)
 - unüberwachter Prozeß: Suche nach Regelmäßigkeiten, Data Mining

Überwachte Klassifizierung von Städten

PIRIT – Einleitung
Problemstellung
Metadaten
explizite Information
Beispiele, Ausblick

► Entscheidungsbaum



Cluster bilden

Bevölkerungsdichte

Lebensqualität

ökonomisches Gewicht

Handelszentren

administrative Einrichtungen

Straßenkreuzungen

Anzahl von Bahnhöfen

(Flug)Häfen, Flüsse, Meer

► Regelmäßigkeiten:

- Hauptstädte an großen Flüssen?
- allg. alle Städte über 100000 EW an großen Flüssen?
- Industriegebiete immer längs von Verkehrsstrassen?
- im Stadtkern größere Gebäude als in Randgebieten?
- in touristischen Gebieten mehr Fahrradwege?
- gewundene Straßen nur in Gebieten mit starken Höhenunterschieden?
- häufig Ackerflächen um Dörfer, danach Wald?
- Verkehrsknoten immer innerhalb von Städten?
- Wohngebiete am Wasser: erhöhter Bodenwert?
- etc.

3 Buzzwords: Metadaten, Semantik, Data Mining

PIRIT – Einleitung

Problemstellung

Metadaten

explizite Information

Beispiele, Ausblick

- ▶ Metadaten – Information über Daten
 - neue **aussagekräftige** Daten über die Daten

- ▶ von „normalen“ Metadaten zu semantisch bedeutenden Metadaten

- Steigerung der Aussagekraft

- ▶ Data Mining – Beispiele

- überwacht, unüberwacht

```
FTSimpleDisplay
<Metadata>
  <Name> Wasserflächen (111) - Objektteil-Linien </Name>
  <Path> M:\heinzle\lckem\Waltrup\arcview\F111_LT.shp </Path>
  <CreateDate> </CreateDate>
  <Keywords>
    Strom, Fluß, Bach
    Graben, Kanal (Wasserwirtschaft)
  </Keywords>
  <Number of Entities> 277 </Number of Entities>
  <ShapeType> Polylines </ShapeType>
  <Min X> 2587421.800 </Min X>
  <Min Y> 5716972.200 </Min Y>
  <Max X> 2594000.000 </Max X>
  <Max Y> 5723128.100 </Max Y>
</Metadata>
```

