Universität Konstanz

„Presentation and Analysis of Spatial Data“

(2) Spatial Data

Thomas Wöhler, Universität Konstanz

Kiev, October, 2016

Agenda

- Types of spatial data
- Terminology
- Data formats
- Data sources:
  - Websites
  - Georeferencing
  - Create your own data

Spatial data

Spatial data reference to a location in the “real” world by means of coordinates:

Latitude	Longitude	Ort	Strasse	Name	Sauberkeit	Bier
47.652502	9.166325	Kreuzlingen	Hauptstr. 34	American Blue Bar	3	5.5

- Informationen about the location
- Informationen about the units of observation (features) and of their characteristics (attributes)

→ Relational database

Types of spatial data

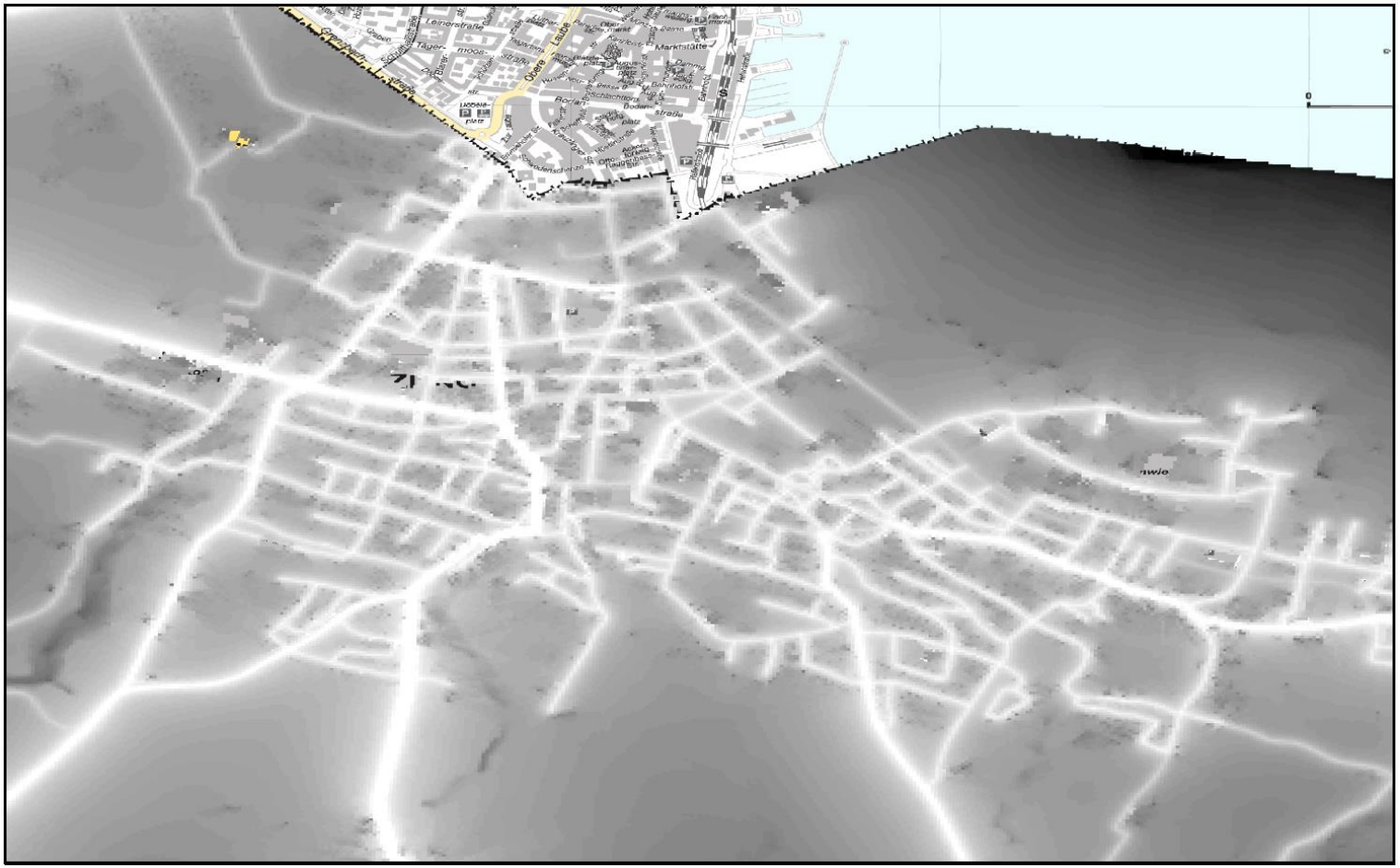
- Raster-data
- Vector-data
  - Point-feature
  - Line-feature
  - Polygon-feature

Raster-data

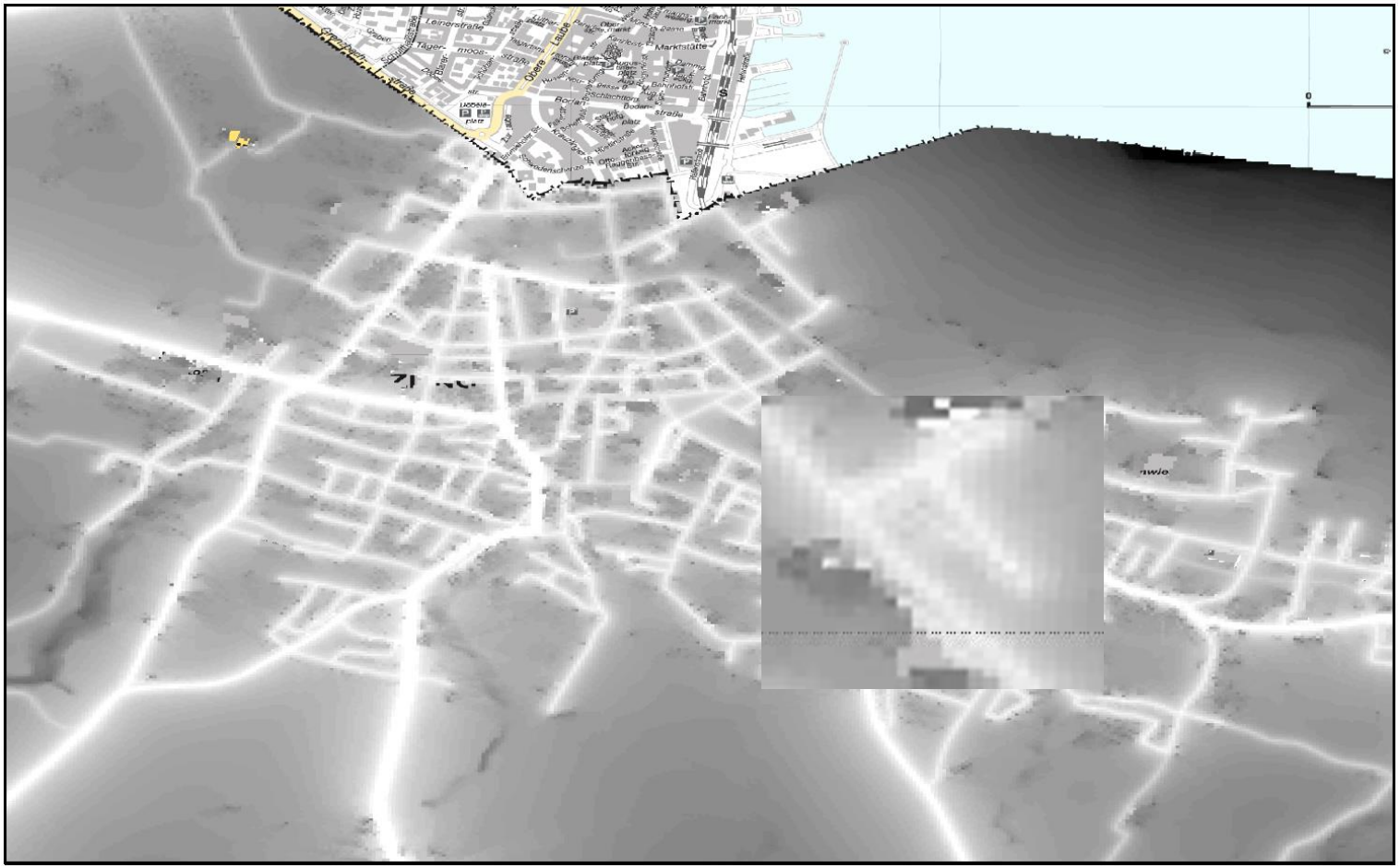
(a) Raster data are often images taken by satellites. Here we can see mountains in the Eastern Cape.

(b) The same raster data, but this time zoomed in. The grid nature of the data can be seen

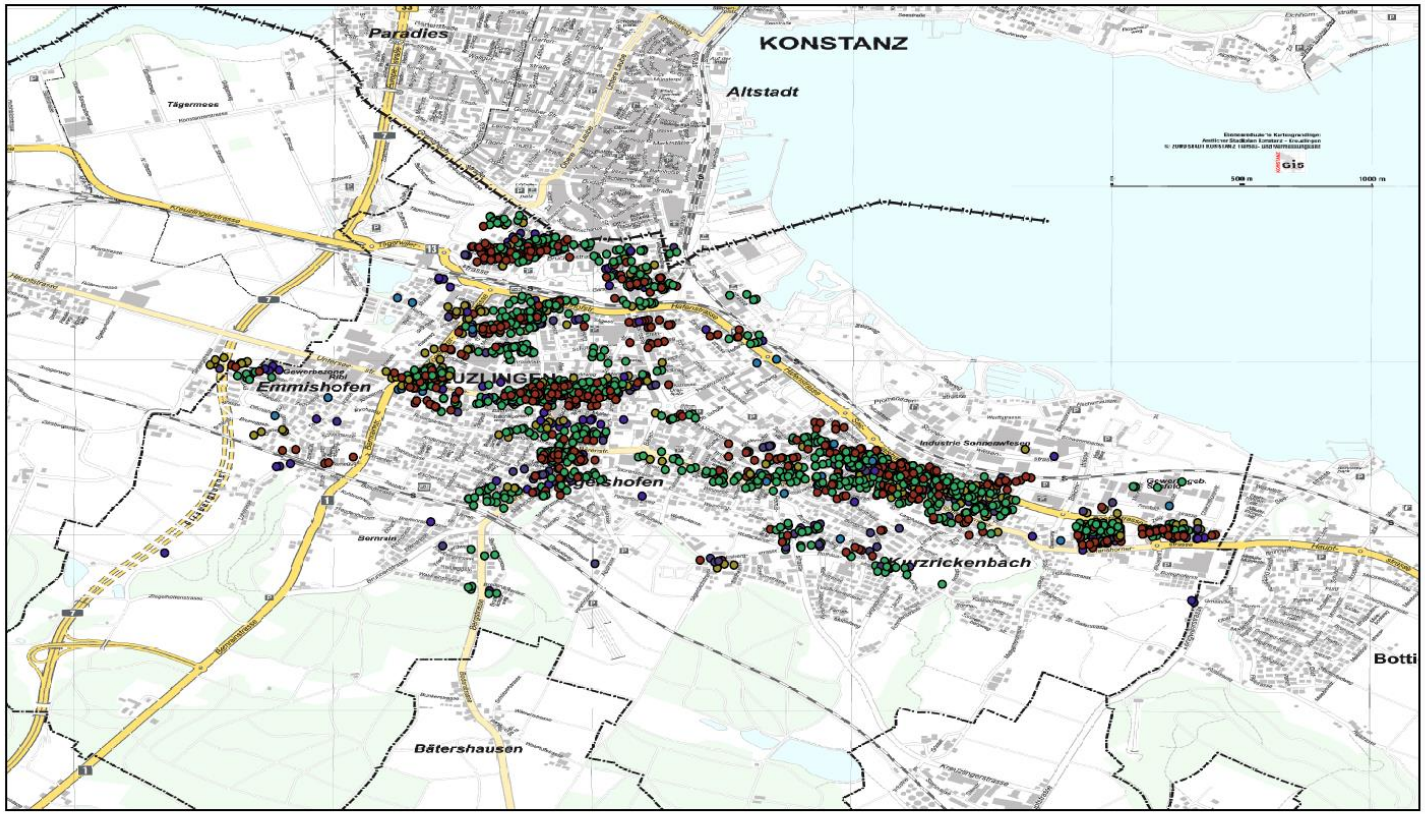
Raster-data: Street noise



Raster-data: Street noise



Vector-data: Point-feature



Einwohner aus einem der Teilstaaten des ehemaligen Jugoslawien und Albanien (2010)

- Bosnien-Herzegovina
- Kosovo / Albanien
- Kroatien
- Mazedonien
- Serbien / Montenegro
- Slowenien

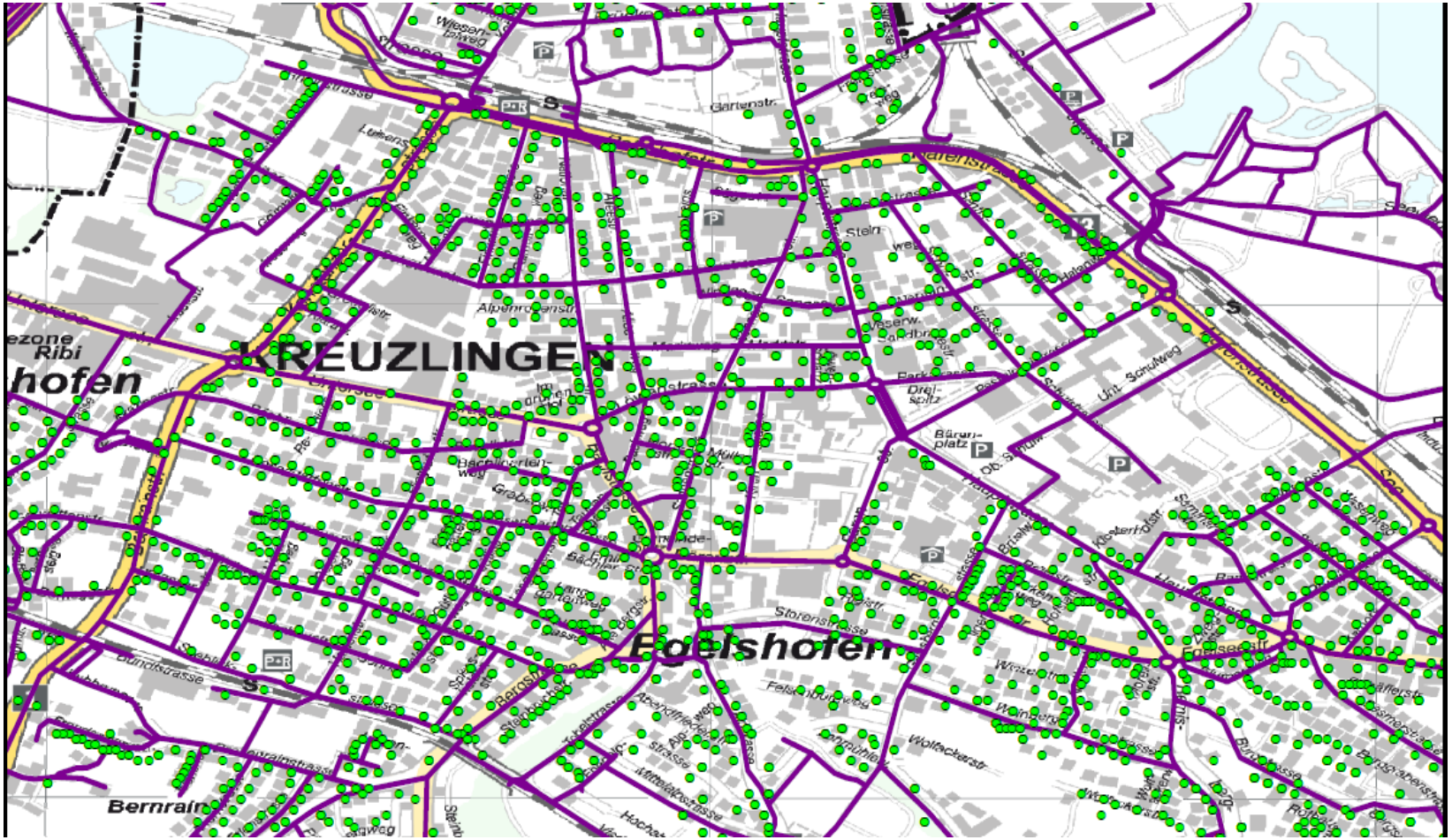
"Digitalisierte Grundrißdaten der amtlichen Flurkarte des Tiefbau- und Vermessungsamtes der Stadt Konstanz Stand 05.07.2010"

Einwohnerdaten der Stadt Kreuzlingen 2010

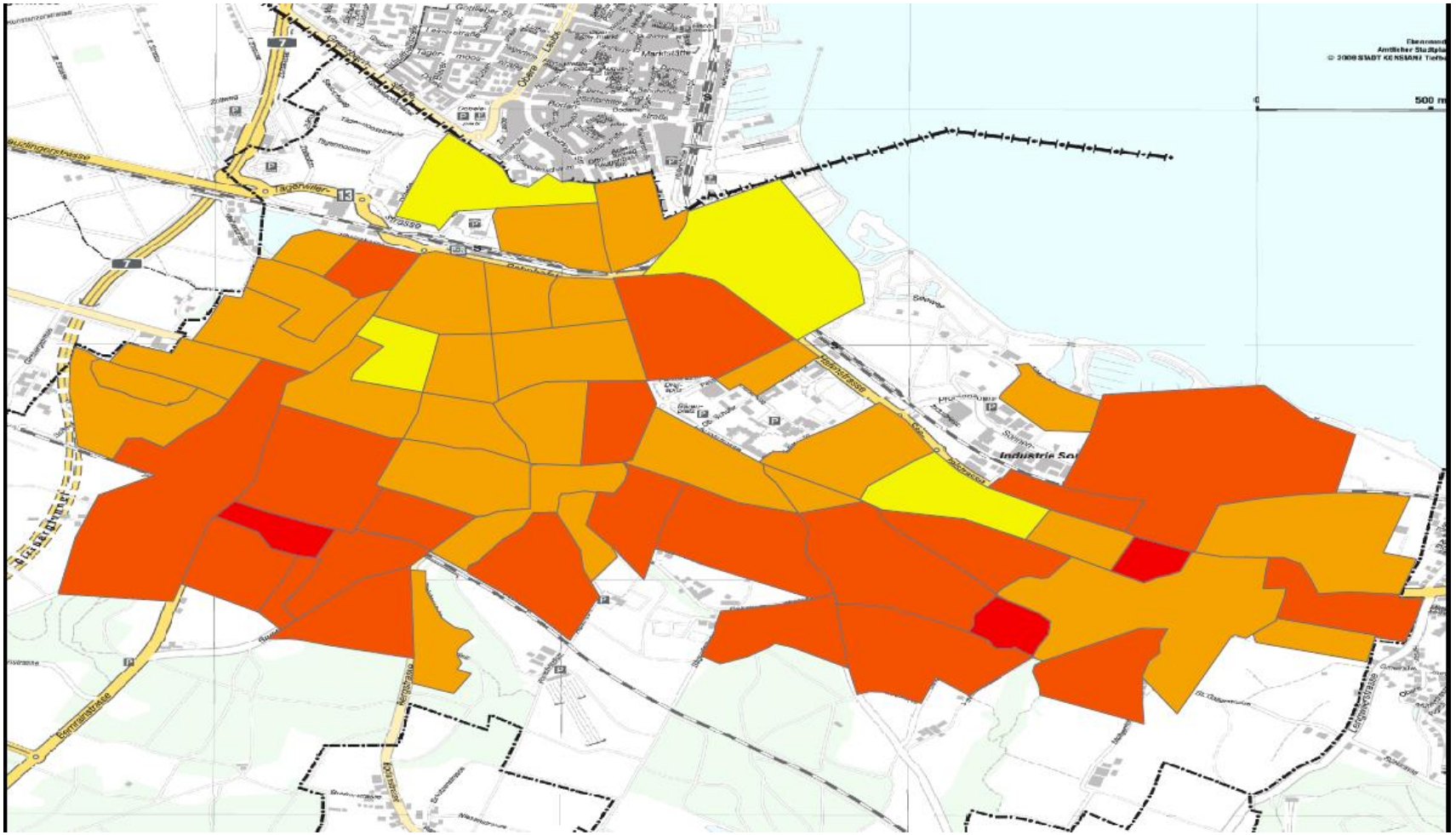
Einzelne Einwohnerpunkte wurden um eine Zufallsdistanz verschoben um Anonymität zu garantieren.

erstellt von Thomas Wöhler (Universität Konstanz)

Vector-data: (Poly-)line-feature

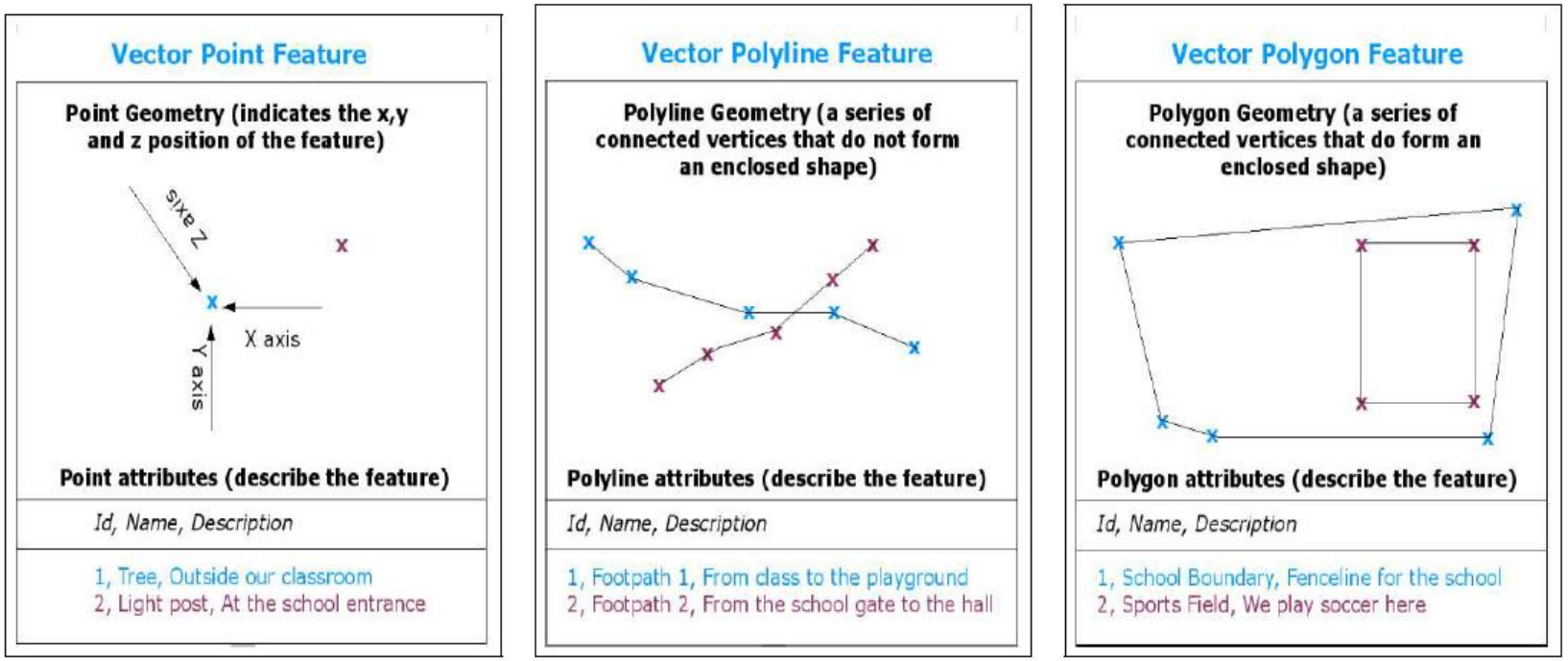


Vector-data: Polygone-feature



Vector-data

Figure 10: Vector point, polyline and polygon geometries



(a) A point feature is described by its X, Y and optionally Z coordinate. The point attributes describe the point e.g. if it is a tree or a lamp post.

(b) A polyline is a sequence of joined vertices. Each vertex has an X, Y (and optionally Z) coordinate. Attributes describe the polyline.

(c) A polygon, like a polyline, is a sequence of vertices. However in a polygon, the first and last vertices are always at the same position.

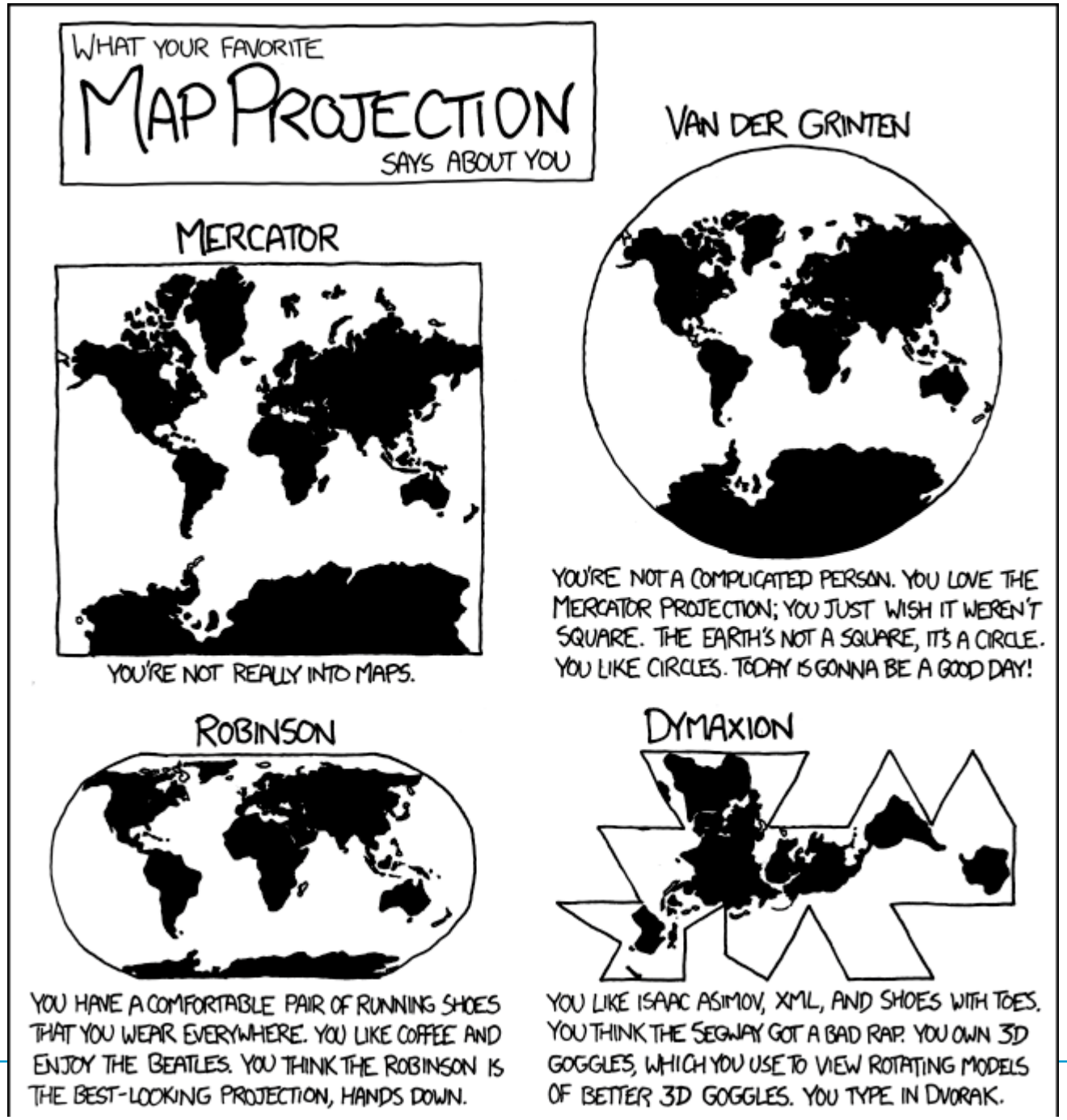
Coordinate systems and Projections

- As many coordinate systems as cartographers (almost).
- Various scales and various developmental histories

- Germany: Deutsches_Hauptdreiecksnetz_Transverse_Mercator,
- Switzerland: CH1903
- Internet: WGS1984

The earth is a sphere!

→ In order to analyse the data (two-dimensional), they have to be projected to two Dimensions.


WHAT YOUR FAVORITE
MAP PROJECTION
SAYS ABOUT YOU
MERCATOR
YOU'RE NOT REALLY INTO MAPS.
VAN DER GRINTEN
YOU'RE NOT A COMPLICATED PERSON. YOU LOVE THE MERCATOR PROJECTION; YOU JUST WISH IT WEREN'T SQUARE. THE EARTH'S NOT A SQUARE, IT'S A CIRCLE. YOU LIKE CIRCLES. TODAY IS GONNA BE A GOOD DAY!
ROBINSON
YOU HAVE A COMFORTABLE PAIR OF RUNNING SHOES THAT YOU WEAR EVERYWHERE. YOU LIKE COFFEE AND ENJOY THE BEATLES. YOU THINK THE ROBINSON IS THE BEST-LOOKING PROJECTION, HANDS DOWN.
DYMAXION
YOU LIKE ISAAC ASIMOV, XML, AND SHOES WITH TOES. YOU THINK THE SEGWAY GOT A BAD RAP. YOU OWN 3D GOGGLES, WHICH YOU USE TO VIEW ROTATING MODELS OF BETTER 3D GOGGLES. YOU TYPE IN DVORAK.

WINKEL-TRIPEL

NATIONAL GEOGRAPHIC ADOPTED THE WINKEL-TRIPEL IN 1998, BUT YOU'VE BEEN A W-T FAN SINCE *LONG* BEFORE "NAT GEO" SHOWED UP. YOU'RE WORRIED IT'S GETTING PLAYED OUT, AND ARE THINKING OF SWITCHING TO THE KAVRAYSKIY. YOU ONCE LEFT A PARTY IN DISGUST WHEN A GUEST SHOWED UP WEARING SHOES WITH TOES. YOUR FAVORITE MUSICAL GENRE IS "POST-".

GOODE HOMOLOSINE

THEY SAY MAPPING THE EARTH ON A 2D SURFACE IS LIKE FLATTENING AN ORANGE PEEL, WHICH SEEMS EASY ENOUGH TO YOU. YOU LIKE EASY SOLUTIONS. YOU THINK WE WOULDN'T HAVE SO MANY PROBLEMS IF WE'D JUST ELECT *NORMAL* PEOPLE TO CONGRESS INSTEAD OF POLITICIANS. YOU THINK AIRLINES SHOULD JUST BUY FOOD FROM THE RESTAURANTS NEAR THE GATES AND SERVE *THAT* ON BOARD. YOU CHANGE YOUR CAR'S OIL, BUT SECRETLY WONDER IF YOU REALLY *NEED* TO.

HOBO-DYER

YOU WANT TO AVOID CULTURAL IMPERIALISM, BUT YOU'VE HEARD BAD THINGS ABOUT GALL-PETERS. YOU'RE CONFLICT-AVERSE AND BUY ORGANIC. YOU USE A RECENTLY-INVENTED SET OF GENDER-NEUTRAL PRONOUNS AND THINK THAT WHAT THE WORLD NEEDS IS A REVOLUTION IN CONSCIOUSNESS.

PLATE CARRÉE
(EQUIRECTANGULAR)

YOU THINK THIS ONE IS FINE. YOU LIKE HOW X AND Y MAP TO LATITUDE AND LONGITUDE. THE OTHER PROJECTIONS OVERCOMPLICATE THINGS. YOU WANT ME TO STOP ASKING ABOUT MAPS SO YOU CAN ENJOY DINNER.

WAGNER BUTTERFLY

A GLOBE!

YES, YOU'RE VERY CLEVER.

PEIRCE QUINCUNCIAL

YOU THINK THAT WHEN WE LOOK AT A MAP, WHAT WE REALLY SEE IS OURSELVES. AFTER YOU FIRST SAW *INCEPTION*, YOU SAT SILENT IN THE THEATER FOR SIX HOURS. IT FREAKS YOU OUT TO REALIZE THAT EVERYONE AROUND YOU HAS A SKELETON INSIDE THEM. YOU *HAVE* REALLY LOOKED AT YOUR HANDS.

WATERMAN BUTTERFLY

REALLY? YOU KNOW THE WATERMAN? HAVE YOU SEEN THE 1909 CAHILL MAP IT'S BASED— ...YOU HAVE A FRAMED REPRODUCTION AT HOME?! WHOA. ...LISTEN, FORGET THESE QUESTIONS. ARE YOU DOING ANYTHING TONIGHT?

GALL-PETERS

I *HATE* YOU.

Coordinate systems and Projections

- When the coordinates are given as latitude and longitude, this points to a geographic (non-projected) coordinate system.
- When the coordinates are given as XY (and stated in meter or kilometer), this points to a projected coordinate system.

- You should transform all layers to a common, projected coordinate system.
- In order to do this, you need a GIS. Mistakes can be detected by using a base map and points with known location.

- Data from the internet (Google, Bing, Openstreetmap) are presented in a Mercator-projection. But you usually get them in latitude und longitude and will have to project them again.

→ EPSG: 4326 (WGS84, geographic coordinate system)
→ EPSG: 3857 (WGS84 Pseudo Mercator, projected coordinate system)

Terminology

Feature	**- Unit with geographic reference, can be points, lines or polygons**
Fields	**- Characteristics of unit (analoguous to variables)**
Attribute	**- Values of the fields**
shape-file	**- Contains the geographic data**
layers	**- Informationen about the presentation of the shape-file in your GIS**

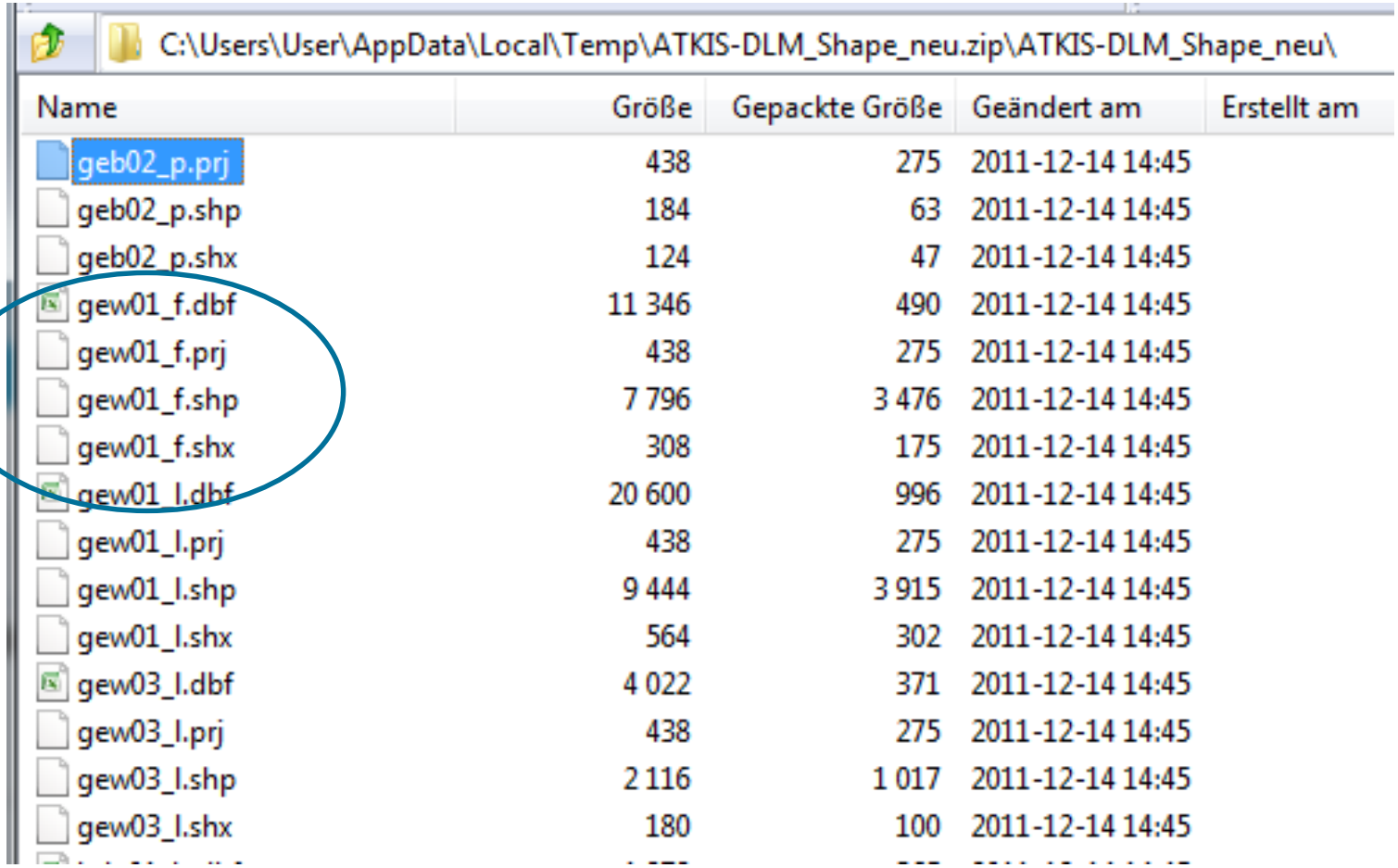
Data format: Shape-files

Proprietary data format of ESRI Corp., de-facto industry-standard

Relational database, cmprising of at least three files:

.shp	geographical information
.dbf	attributes in dBASE-Format
.shx	Index to combine geographical information and attributes
+	.prj Projection of the data

Data format: Shape-files



Name	Größe	Gepackte Größe	Geändert am	Erstellt am
geb02_p.prj	438	275	2011-12-14 14:45	
geb02_p.shp	184	63	2011-12-14 14:45	
geb02_p.shx	124	47	2011-12-14 14:45	
gew01_f.dbf	11 346	490	2011-12-14 14:45	
gew01_f.prj	438	275	2011-12-14 14:45	
gew01_f.shp	7 796	3 476	2011-12-14 14:45	
gew01_f.shx	308	175	2011-12-14 14:45	
gew01_l.dbf	20 600	996	2011-12-14 14:45	
gew01_l.prj	438	275	2011-12-14 14:45	
gew01_l.shp	9 444	3 915	2011-12-14 14:45	
gew01_l.shx	564	302	2011-12-14 14:45	
gew03_l.dbf	4 022	371	2011-12-14 14:45	
gew03_l.prj	438	275	2011-12-14 14:45	
gew03_l.shp	2 116	1 017	2011-12-14 14:45	
gew03_l.shx	180	100	2011-12-14 14:45	

Data format: csv

- Comma-seperated values in a text file
- Commonly used format for transfer of data between various computer programs
- Tables
- XY tables (point data)

Data sources: Opendata

- Official administrative units in Europe (NUTS): Eurostat
- Data of Openstreetmap (http://download.geofabrik.de/)
- https://www.census.gov / http://www.zensus2011.de
- Ask your local administration

Data sources: commercial data

- Data of Openstreetmap (http://download.geofabrik.de/)
- Google Maps / Bing Maps / Nokia Here
- In Germany: microm, gfk - market research companies

Data sources: Georeference addresses

- You can georeference postal addresses, i.e. assign latitude and longitude to observational units based on a postal address.
- This is done by using APIs automatically via Stata / R / Python
- Development very dynamic and access changes frequently.
- Available data:
    - Google Maps (very good data quality, 2500 queries a day, after that you have to pay 50ct / 1000 addresses (or use another IP-address), flexible, Google saves addresses and the usage does not comply with „Terms and Regulations“)
    - Bing (similar to Google)
    - Nokia Here (best data quality)
    - Openstreetmaps (free data, data quality unknown, access more difficult)

Data sources: Georeference addresses

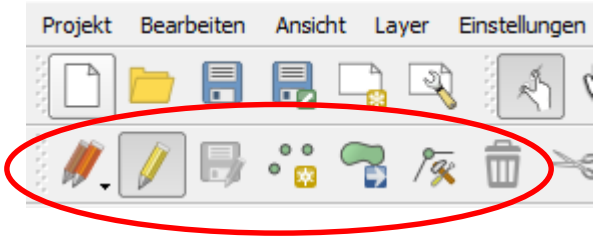
- For getting started: www.gpsvisualizer.com/geocoder/
- Organisation of the addresses in a csv-file
- Cyrillic?

Steps:

1. Georeferencing of addresses
2. Keep order of addresses
3. Generate identifiers
4. Match with other characteristics

Data sources: Create your own data

- Create a new shape-file with QGIS (→ „Layer“ → „Create Layer“)
- Choose the correct type (Point, Line, Polygon) and the appropriate coordinate system
- Choose a basemap (→ „Web“ → „Openlayers Plugin“)
- Edit:



- Or: Drag, add or delete points in a existing shape-file
- Add or change attributes
- Save by leaving editing mode

Universität Konstanz

Thank You
For Your Attention!

Thomas Wöhler
Fachbereich Geschichte und Soziologie, AG Hinz

Tel.: +49 (0) 75 31/88 - 3345
thomas.woehler@uni-konstanz.de