

UNIVERSITY OF TWENTE.

Lesson 5: DATA ACQUISITION - GIS
Textbook 8.2 & 8.3

RICHARD KNIPPERS

ivc FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION 2023

DIRECT SPATIAL DATA ACQUISITION

- Data directly captured from the environment
- Primary data

Terrestrial survey

Field surveys

Aerial surveys and satellite remote sensing

GPS / Mobile GIS

CONTENTS

- Data acquisition
 - Direct and indirect data acquisition
 - Geoportals & webservices
 - Volunteered Geographic Information (VGI)
- Digitizing from existing documents
 - Digitizing
 - Scanning
 - Vectorization
- Data preparation
 - Data checks and repairs

ivc UNIVERSITY OF TWENTE. 2

INDIRECT SPATIAL DATA ACQUISITION

- Data derived from existing sources
- Data that have been collected for other purposes
- Secondary data

Geoportals

Digitizing

ivc UNIVERSITY OF TWENTE. 4

1

2

GEOPORTALS

- Spatial data can be obtained via a **Geoportal**. They enable the discovery and delivery of data through a service-based architecture known as **Spatial Data Infrastructure**¹.

e.g. Nationaal Georegister: <http://nationaalgeoregister.nl>

ivc UNIVERSITY OF TWENTE. 3

¹ SDI will be discussed in greater detail in the next course.

GIS WEB SERVICES

- A Web Map Service (WMS) is a web server that provides access to raster data, a Web Feature Service (WFS) is a web server that provides access to vector (GML) data, and a Web Coverage Service (WCS) is a web server that provides access to imagery.
- These web service standards for spatial data were developed by the Open Geospatial Consortium (OGC). Some of the most recognized services in this context are WMS, WFS, and WCS <https://www.ogc.org>

ivc UNIVERSITY OF TWENTE. 7

DUTCH NATIONAL GEOPORTAL

<http://nationaalgeoregister.nl>

ivc National Georegister

WMS service URL: <https://www.geoportaloverijssel.nl/geoserver/ows?service=WFS&version=1.0.0&request=GetFeature>

WFS service URL: <https://www.geoportaloverijssel.nl/geoserver/ows?service=WFS&version=1.0.0&request=GetFeature>

WEB MAP SERVICE OPERATION

Client

Web Server

1. Creates Request

2. Request sent to web service

3. Receives request

4. Calls the map server to request map

5. Map Server generates map as image file

6. Creates response with image file

7. Response sent to client

8. Receives response, parses to extract image and updates map

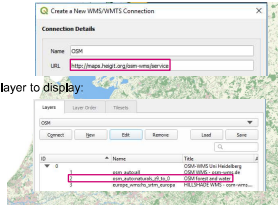
ivc UNIVERSITY OF TWENTE. 4

3

4

WEB MAP SERVICE OPERATION IN QGIS

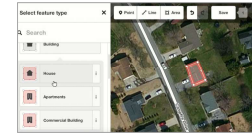
- Service URL to load OpenStreetMap data from a WMS service:
<http://maps.heigit.org/osm-wms/service?REQUEST=GetCapabilities&SERVICE=WMS>
- Create connection in QGIS (*Data Source Manager>WMS/WMST*):



- Select WMS data layer to display:

VOLUNTEERED GEOGRAPHIC INFORMATION (VGI) CROWDSOURCED DATA COLLECTION

- Citizens can play a complementary role by providing geo-referenced information known as Volunteered Geographic Information (VGI).
- Products such as Wikimapia and OpenStreetMap (OSM) can be considered as VGI. Citizens can directly participate in the collection of spatial data. It is made by people like you and me.



Adding a building to OpenStreetMap

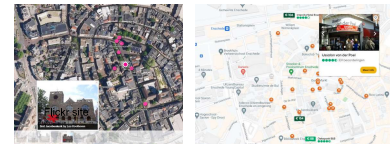
https://www.youtube.com/watch?v=E1YJV6I_rhY

CONTENTS

- Data acquisition
 - Direct and indirect data acquisition
 - Geoportals & webservices
 - Volunteered Geographic Information (VGI)
- Digitizing from existing documents
 - Digitizing
 - Scanning
 - Vectorization
- Data preparation
 - Data checks and repairs

VOLUNTEERED GEOGRAPHIC INFORMATION (VGI)

- A typical example of VGI is when users add their locations and information about themselves on a map. Therefore, sites such as TripAdvisor, Flickr, Twitter and Instagram can also be considered as VGI.

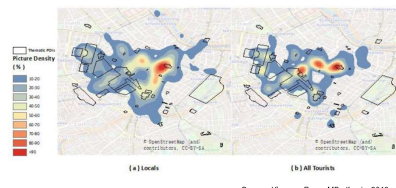


Flickr

TripAdvisor

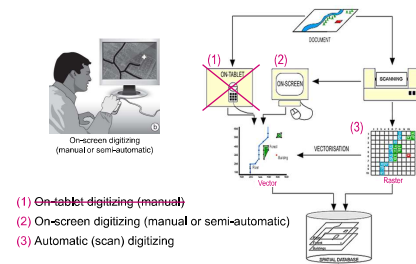
FOOTPRINTS OF TOURISTS DERIVED FROM FLICKR DATA

- Footprints of locals and tourists in the city centre of Vienna derived from Flickr data.



Source: Yingwen Deng, MSc Thesis, 2019

DIGITIZING TECHNIQUES

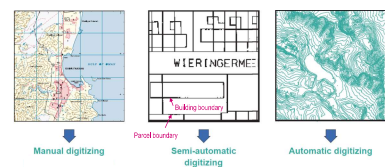


CONTENTS

- Data acquisition
 - Direct and indirect data acquisition
 - Geoportals & webservices
 - Volunteered Geographic Information (VGI)
- Digitizing from existing documents
 - Digitizing
 - Scanning
 - Vectorization
- Data preparation
 - Data checks and repairs

SELECTING A DIGITIZING METHOD

- Manual digitizing > complex documents / interpretations from satellite imagery or aerial photographs
- Semi-automatic digitizing > simple documents that require some interpretation
- Automatic digitizing > simple documents with one type of information



MAP REGISTRATION PROCESS

Reference point 1

reference point 2

reference point 3

reference point 4

RMSE

Auto Adjust Transformation: 1st Order Polynomial (affine) Total RMS Error: 0.0035

VECTOR DATA FORMAT
GEOPACKAGE

- A **geopackage** stores point, line, and polygon data in a relational database as tables.

Enschede.gpkg → Geopackage

Buildings

MunBoundary

Railways

Stations

Tables

Field for representing spatial features

id	geom	TrackID	Present
1	Polygon	17	1
2	Polygon	21	1
3	Polygon	24	1
4	Polygon	25	1
5	Polygon	26	1
6	Polygon	1	0
7	Polygon	2	0

VECTOR DATA FORMAT
SHAPEFILE

- A **Shapefile** is vector data structure for storing spatial and attribute data
- It contains only one geometry type (either point, line or polygon)
- It consists of a minimum of three files: **shapefile.shp**, **shapefile.shx** and **shapefile.dbf**
- It is a non-topological data structure (polygon-by-polygon)

Stations_Hengelo_Enschede.shp

Stations_Hengelo_Enschede DBF File

Stations_Hengelo_Enschede PRJ File

Stations_Hengelo_Enschede SHP File

Stations_Hengelo_Enschede.shx SHX File

DIGITIZING MODES
POINT MODE AND STREAM MODE

- When you trace a line feature you can do this in point mode or stream mode. In point mode you digitize each point of the line. In stream mode you follow the line with your cursor and the vertices are automatically recorded depending on a preset tolerance from a straight line.

Point mode digitizing

Stream mode digitizing

Detected point

Tolerance

manually selected points

recorded point

selected point

recorded line

selected line

Advanced Digitizing Toolbar

DIGITIZING TIPS

- Use snapping to avoid overshoots and undershoots
- Avoid duplicate lines (adjoining polygons)

Overshoot

Undershoot

Sliver

Duplicate line

DIGITIZING TIPS
ASSOCIATING ATTRIBUTES

Identical ID number!

Missing ID number!

ObjectID: 15 Class: Building Area: 7623 Total value: 52100

ObjectID: 14 Class: secondary road Area: 10223 Total value: 34256

ObjectID: 8 Class: Wood Area: 10521 Total value: 24500

ObjectID: 6 Class: secondary road in tunnel Area: 185 Total value: 34526

DIGITIZING TIPS
DIGITIZING AREAS

- In certain cases it is more convenient to digitize area objects as lines and thereafter use an automatic procedure to construct the polygons.

No polygon created

Automatic polygon formation

DIGITIZING TIPS
SNAPPING TOOL IN QGIS

Enable Snapping

Snapping options

Vertex

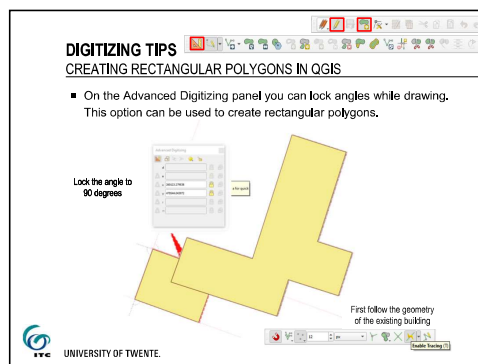
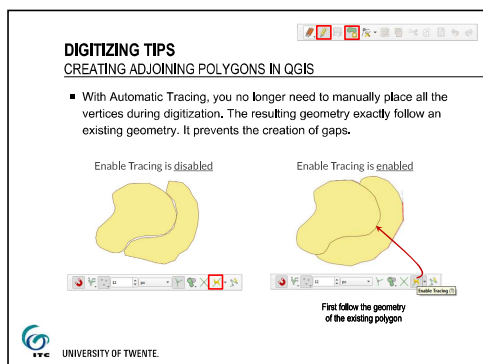
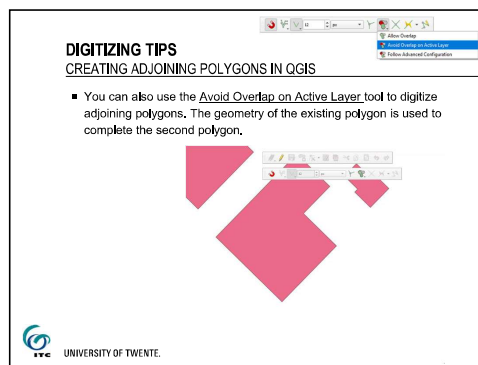
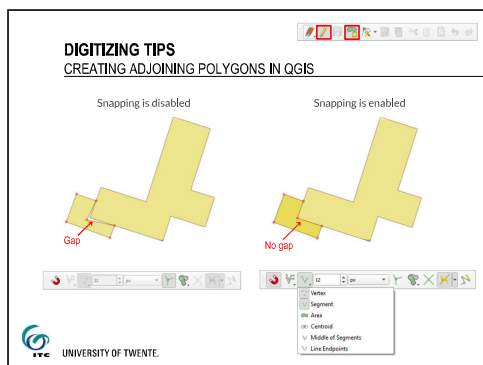
Segment

Area

Centroid

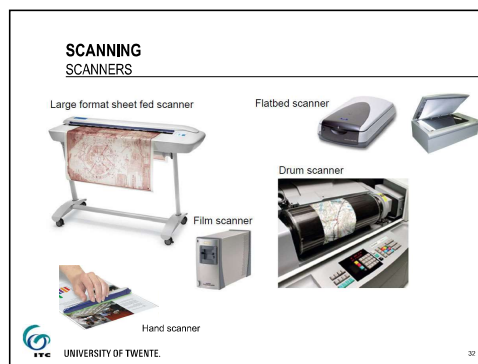
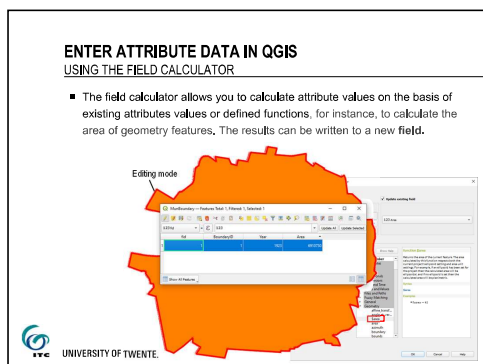
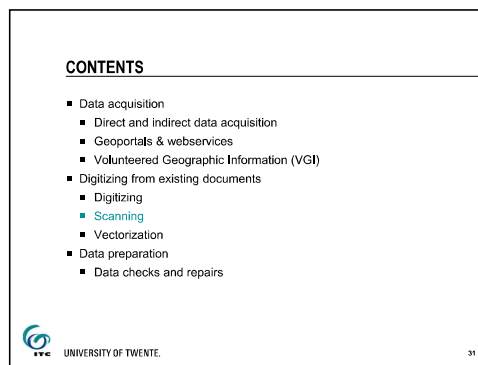
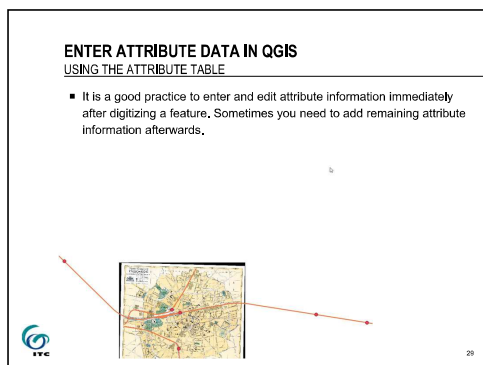
Middle of Segments

Line Endpoints



13

14



15

16

SCANNING

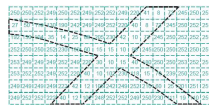
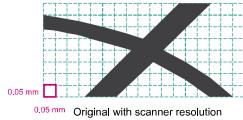
MODES AND RESOLUTION

Scanning resolution:

- millimeters (e.g. 0,05 mm)
- microns (e.g. 50 µm)
- dots per inch (e.g. 500 dpi)
- nr. of pixels per inch (2,54 cm)

Scanning modes:

- Bitmap (black=1 white=0)
- Grey-scale (256 grey values per pixel)
- Color (Red, Green, Blue)



CONTENTS

- Data acquisition
 - Direct and indirect data acquisition
 - Geoportals & webservices
 - Volunteered Geographic Information (VGI)
- Digitizing from existing documents
 - Digitizing
 - Scanning
 - Vectorization
- Data preparation
 - Data checks and repairs

SCANNING

SELECTING SCANNER RESOLUTION

Manual digitizing:

- Printed map: 200-300 dpi
- Aerial photographs: ≥ 800 dpi

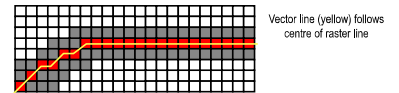
Semi-automatic / automatic digitizing:

- Printed map: 300-600 dpi (depending on the thickness of the thinnest lines)



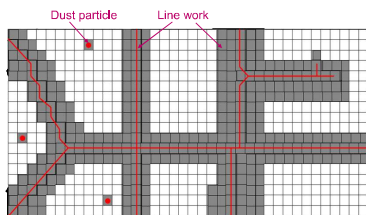
VECTORIZATION

DISPLACEMENT OF VECTOR LINES



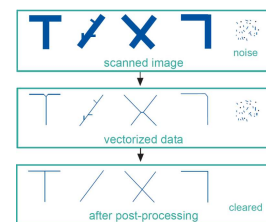
VECTORIZATION

PROCESS: SKELETONIZING



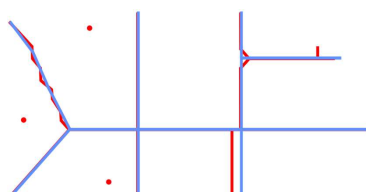
VECTORIZATION

PROCESS: POST-PROCESSING



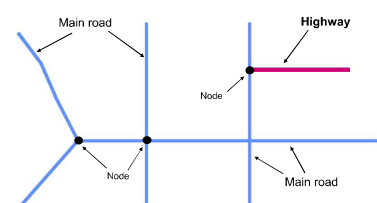
VECTORIZATION

PROCESS: POST-PROCESSING



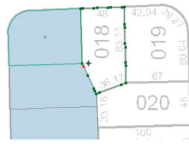
VECTORIZATION

PROCESS: FEATURE FORMING



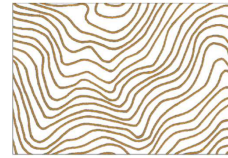
SEMI-AUTOMATIC DIGITIZING RASTER TRACING

- Semi-automatic digitizing, also known as raster tracing, involves the process of vectorizing line and polygon features through a semi-automated approach, where line features are traced partially using automation.



AUTOMATIC DIGITIZING

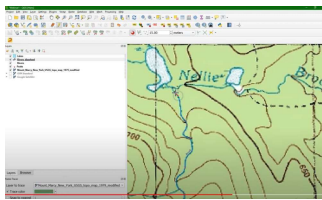
- Automatic digitizing is the process of vectorizing line and polygon features through a completely automated approach, wherein line features are traced entirely using automation.



SEMI-AUTOMATIC DIGITIZING RASTER TRACER



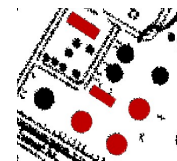
- In QGIS, there is a plugin called Raster Tracer. The Raster Tracer can be used for semi-automatic digitizing of an underlying raster layer.



<https://www.youtube.com/watch?v=D3Mh4wpSEMM>

AUTOMATIC DIGITIZING SHAPE RECOGNITION

- You can also capture vector features of a certain shape, such as a building or storage tank, via the shape recognition tools.



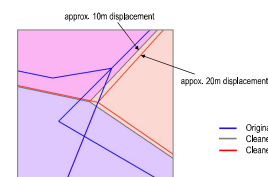
Click the raster feature to capture buildings and storage tanks

CONTENTS

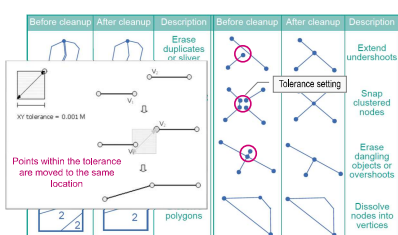
- Data acquisition
 - Direct and indirect data acquisition
 - Geoportals & webservices
 - Volunteered Geographic Information (VGI)
- Digitizing from existing documents
 - Digitizing
 - Scanning
 - Vectorization
- Data preparation
 - Data checks and repairs

CLEAN-UP OPERATIONS FOR VECTOR DATA TOLERANCE SETTING

- Settings such as a tolerance for the cleaning of overshoots, undershoots or node clusters have to be set carefully to avoid that lines displace, do not connect or connect where they should not connect.

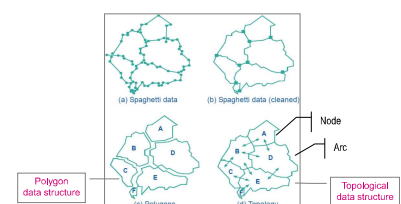


CLEAN-UP OPERATIONS FOR VECTOR DATA



CLEAN-UP OPERATIONS FOR AREAS

- The topological data structure can aid in data editing area features.



QGIS CLEAN-UP OPERATIONS

TOPOLOGY EDITING

Enable topological editing

- Topological editing in QGIS can be used to edit the shared edges of two adjacent polygons

Standard editing

Topology editing

Gap

No Gap

UNIVERSITY OF TWENTE

QGIS CLEAN-UP OPERATIONS

TOPOLOGY EDITING

- The **Geometry Checker** plugin can also be used to validate the geometry of data layers. Any identified errors can be **automatically** corrected.

Layer	Object ID	Error	Coordinates	Value
buildings_checkbuildings	43485	Duplicate	252627.52, 475569.03	buildings_checkBuild-19222

UNIVERSITY OF TWENTE

QGIS CLEAN-UP OPERATIONS

TOPOLOGY EDITING

- The **Topology Checker** plugin can be used to check the geometry validity by evaluating a set of topology rules. The identified errors require **manual** editing.

1. Define topology rules

2. Errors found are colored red

UNIVERSITY OF TWENTE

QGIS CLEAN-UP OPERATIONS

TOPOLOGY EDITING

- The **Geometry Checker** plugin can also be used to validate the geometry of data layers. Any identified errors can be **automatically** corrected.

INPUT

UNIVERSITY OF TWENTE

25

26

CLEAN-UP OPERATIONS

WORKED-OUT EXAMPLE IN QGIS

Layer	Object ID	Error	Coordinates	Value
buildings	1	Self-intersecting line	252627.52, 475569.03	buildings
buildings	2	Self-intersecting line	252627.52, 475569.03	buildings
buildings	3	Self-intersecting line	252627.52, 475569.03	buildings
buildings	4	Self-intersecting line	252627.52, 475569.03	buildings
buildings	5	Self-intersecting line	252627.52, 475569.03	buildings
buildings	6	Self-intersecting line	252627.52, 475569.03	buildings
buildings	7	Self-intersecting line	252627.52, 475569.03	buildings
buildings	8	Self-intersecting line	252627.52, 475569.03	buildings
buildings	9	Self-intersecting line	252627.52, 475569.03	buildings
buildings	10	Self-intersecting line	252627.52, 475569.03	buildings
buildings	11	Self-intersecting line	252627.52, 475569.03	buildings
buildings	12	Self-intersecting line	252627.52, 475569.03	buildings
buildings	13	Self-intersecting line	252627.52, 475569.03	buildings
buildings	14	Self-intersecting line	252627.52, 475569.03	buildings
buildings	15	Self-intersecting line	252627.52, 475569.03	buildings
buildings	16	Self-intersecting line	252627.52, 475569.03	buildings
buildings	17	Self-intersecting line	252627.52, 475569.03	buildings
buildings	18	Self-intersecting line	252627.52, 475569.03	buildings
buildings	19	Self-intersecting line	252627.52, 475569.03	buildings
buildings	20	Self-intersecting line	252627.52, 475569.03	buildings
buildings	21	Self-intersecting line	252627.52, 475569.03	buildings
buildings	22	Self-intersecting line	252627.52, 475569.03	buildings
buildings	23	Self-intersecting line	252627.52, 475569.03	buildings
buildings	24	Self-intersecting line	252627.52, 475569.03	buildings
buildings	25	Self-intersecting line	252627.52, 475569.03	buildings
buildings	26	Self-intersecting line	252627.52, 475569.03	buildings
buildings	27	Self-intersecting line	252627.52, 475569.03	buildings
buildings	28	Self-intersecting line	252627.52, 475569.03	buildings
buildings	29	Self-intersecting line	252627.52, 475569.03	buildings
buildings	30	Self-intersecting line	252627.52, 475569.03	buildings
buildings	31	Self-intersecting line	252627.52, 475569.03	buildings
buildings	32	Self-intersecting line	252627.52, 475569.03	buildings
buildings	33	Self-intersecting line	252627.52, 475569.03	buildings
buildings	34	Self-intersecting line	252627.52, 475569.03	buildings
buildings	35	Self-intersecting line	252627.52, 475569.03	buildings
buildings	36	Self-intersecting line	252627.52, 475569.03	buildings
buildings	37	Self-intersecting line	252627.52, 475569.03	buildings
buildings	38	Self-intersecting line	252627.52, 475569.03	buildings
buildings	39	Self-intersecting line	252627.52, 475569.03	buildings
buildings	40	Self-intersecting line	252627.52, 475569.03	buildings
buildings	41	Self-intersecting line	252627.52, 475569.03	buildings
buildings	42	Self-intersecting line	252627.52, 475569.03	buildings
buildings	43	Self-intersecting line	252627.52, 475569.03	buildings
buildings	44	Self-intersecting line	252627.52, 475569.03	buildings
buildings	45	Self-intersecting line	252627.52, 475569.03	buildings
buildings	46	Self-intersecting line	252627.52, 475569.03	buildings
buildings	47	Self-intersecting line	252627.52, 475569.03	buildings
buildings	48	Self-intersecting line	252627.52, 475569.03	buildings
buildings	49	Self-intersecting line	252627.52, 475569.03	buildings
buildings	50	Self-intersecting line	252627.52, 475569.03	buildings
buildings	51	Self-intersecting line	252627.52, 475569.03	buildings
buildings	52	Self-intersecting line	252627.52, 475569.03	buildings
buildings	53	Self-intersecting line	252627.52, 475569.03	buildings
buildings	54	Self-intersecting line	252627.52, 475569.03	buildings
buildings	55	Self-intersecting line	252627.52, 475569.03	buildings
buildings	56	Self-intersecting line	252627.52, 475569.03	buildings
buildings	57	Self-intersecting line	252627.52, 475569.03	buildings
buildings	58	Self-intersecting line	252627.52, 475569.03	buildings
buildings	59	Self-intersecting line	252627.52, 475569.03	buildings
buildings	60	Self-intersecting line	252627.52, 475569.03	buildings
buildings	61	Self-intersecting line	252627.52, 475569.03	buildings
buildings	62	Self-intersecting line	252627.52, 475569.03	buildings
buildings	63	Self-intersecting line	252627.52, 475569.03	buildings
buildings	64	Self-intersecting line	252627.52, 475569.03	buildings
buildings	65	Self-intersecting line	252627.52, 475569.03	buildings
buildings	66	Self-intersecting line	252627.52, 475569.03	buildings
buildings	67	Self-intersecting line	252627.52, 475569.03	buildings
buildings	68	Self-intersecting line	252627.52, 475569.03	buildings
buildings	69	Self-intersecting line	252627.52, 475569.03	buildings
buildings	70	Self-intersecting line	252627.52, 475569.03	buildings
buildings	71	Self-intersecting line	252627.52, 475569.03	buildings
buildings	72	Self-intersecting line	252627.52, 475569.03	buildings
buildings	73	Self-intersecting line	252627.52, 475569.03	buildings
buildings	74	Self-intersecting line	252627.52, 475569.03	buildings
buildings	75	Self-intersecting line	252627.52, 475569.03	buildings
buildings	76	Self-intersecting line	252627.52, 475569.03	buildings
buildings	77	Self-intersecting line	252627.52, 475569.03	buildings
buildings	78	Self-intersecting line	252627.52, 475569.03	buildings
buildings	79	Self-intersecting line	252627.52, 475569.03	buildings
buildings	80	Self-intersecting line	252627.52, 475569.03	buildings
buildings	81	Self-intersecting line	252627.52, 475569.03	buildings
buildings	82	Self-intersecting line	252627.52, 475569.03	buildings
buildings	83	Self-intersecting line	252627.52, 475569.03	buildings
buildings	84	Self-intersecting line	252627.52, 475569.03	buildings
buildings	85	Self-intersecting line	252627.52, 475569.03	buildings
buildings	86	Self-intersecting line	252627.52, 475569.03	buildings
buildings	87	Self-intersecting line	252627.52, 475569.03	buildings
buildings	88	Self-intersecting line	252627.52, 475569.03	buildings
buildings	89	Self-intersecting line	252627.52, 475569.03	buildings
buildings	90	Self-intersecting line	252627.52, 475569.03	buildings
buildings	91	Self-intersecting line	252627.52, 475569.03	buildings
buildings	92	Self-intersecting line	252627.52, 475569.03	buildings
buildings	93	Self-intersecting line	252627.52, 475569.03	buildings
buildings	94	Self-intersecting line	252627.52, 475569.03	buildings
buildings	95	Self-intersecting line	252627.52, 475569.03	buildings
buildings	96	Self-intersecting line	252627.52, 475569.03	buildings
buildings	97	Self-intersecting line	252627.52, 475569.03	buildings
buildings	98	Self-intersecting line	252627.52, 475569.03	buildings
buildings	99	Self-intersecting line	252627.52, 475569.03	buildings
buildings	100	Self-intersecting line	252627.52, 475569.03	buildings

UNIVERSITY OF TWENTE

CLEAN-UP OPERATIONS

WORKED-OUT EXAMPLE IN QGIS

Post your questions in the discussion board!

UNIVERSITY OF TWENTE

VIDEO DEMONSTRATIONS AND EXERCISE

- Video demonstrations (Canvas):
 - Practical aspects of digitizing and editing
 - How to digitize and edit line and area features
- Exercise: Data entry
 - Add new features (points, lines, polygons) by digitizing
 - Convert spreadsheet data to a spatial layer
 - Enter attribute data
 - Validate and edit data using topology rules
 - Search data in a Geoportal
 - Optional activity: Vector data cleanup operation

UNIVERSITY OF TWENTE

27

28