ℹ **Vector topology in Context**

> *"Topology expresses the spatial relationships between connecting or adjacent vector features (points, polylines, and polygons) in a GIS. "*

Vector data has to conform to the topological model e.g. two lines in a roads vector layer that intersect should be joined by the same node. Topology is necessary for carrying out some types of spatial analysis, such as network analysis." In this module, we will inspect vector data for topology errors and explore ways in which to clean the dataset so that it conforms to the vector topology model.

The vector topology model stipulates that:

- Boundaries should not cross each other (i.e. boundaries which would cross must be split at their intersection to form distinct boundaries). Lines can cross each other e.g. bridges over rivers.
- Lines and boundaries share nodes only if their endpoints are identical.
- Common area boundaries should appear only once.
- Areas must be explicitly closed.
- It is recommended that lines, points and polygon should be stored separately.

You try:

Goal: To learn the basics of maintaining vector topology and how to clean data so that it adheres to the vector topology model.

- Load the raster layer and vector layer into QGIS. Ensure that the vector layer is rendered on top of the raster.
- Symbolize the vector layer using the roads_style specified. What can you infer about the layer ?.
- Open the properties of the vector layer and choose *layer rendering* under the *Symbology tab*.
- On *blending* mode select *feature blending* and choose *multiply* from the drop-down.
- Zoom to a scale of 2500. What do you see from the roads ?.
- Open the attributes and note the number of features that are available from the layer.
- In the *Processing* Menu select toolbox and search for the algorithm *Delete Duplicates geometries* and run it.

- Open the attributes of the new layer generated above and note the number of features available. How many features have been deleted from the roads_unclean layer?

- Use the **topology checker** to validate against any topology errors from the layer that has no duplicates (clean).
  What kind of errors do you see for the road? How can we clean up the layer ?**

- Reproject the road layer (clean) without duplicates to the specified **CRS.**

- In the processing toolbox use v.clean algorithm and choose the *snap* tool. Select appropriate distances to use.

NB: Ideally start by using a very big number like 100 and note the difference. Then start again this time starting with small numbers with small increments.

- Investigate other plugins or tools native to QGIS which can be used to clean the vector layer.
- Repeat this procedure with another layer preferably a polygon and note the different types of tools that are available for the dataset.

Name	Expectation
Background Layer	Google Satellite or topology_cleaning/ 2821AD_21_2014_747_RGB_RECT.tif
data	roads_unclean
Validity tools (Plugins)	Topology checker, Geometry validity
CRS	EPSG:3857
Cleaning tools	V.clean
Roads Style	Simple line, Colour - #e31a1c, Stroke width - 0.86

More about

The vector topology model stipulates that:

- Boundaries should not cross each other (i.e. boundaries which would cross must be split at their intersection to form distinct boundaries). Lines can cross each other e.g. bridges over rivers.
- Lines and boundaries share nodes only if their endpoints are identical.
- Common area boundaries should appear only once.
- Areas must be explicitly closed.
- It is recommended that lines, points and polygon should be stored separately.

Data that does not adhere to the topological model inhibits spatial analysis to be done and will always give erroneous results. Typical errors are:

- Undershoots - Occurs when digitised roads do not intersect.
- Overshoots - Occurs if a line ends beyond the line it should.
- Slivers - Occurs when the vertices of two or more polygons do not match up on their borders.

GRASS GIS is a fully topological GIS package and using the GRASS algorithms to check and clean data helps in generating clean vector data which can be used in an analysis.

✅ Check your knowledge:

1. Why is vector topology important:
    a. *It enhances spatial analysis because distances across a road network can be accurate*
    b. *It ensures that the symbology of the layers is accurate*
    c. *It allows vector data to be converted into different formats i.e. KML*

2. Which type of data should be checked for topology errors:
    a. *A satellite image that contains roads*
    b. *All layers in QGIS should checked*
    c. *A GIS vector dataset where each polygon represents a type of building*

3. When cleaning polygon is it important to activate snapping:
    a. *True*
    b. *False*

📋 Further reading:

- Plugins_topology_checker: http://docs.qgis.org/2.14/pl/docs/user_manual/plugins/plugins_topology_checker.html
- Topology: https://docs.qgis.org/2.14/en/docs/gentle_gis_introduction/topology.html

Download the sample data for the lesson from http://changelog.inasafe.org/media/images/lesson/worksheet/external_data/e3d74596f51c26a98abe017d37c481251b2ace35.zip.