HeiGIT | HEIDELBERG INSTITUTE FOR GEOINFORMATION TECHNOLOGY

ohsome FOR DATA QUALITY ASSESSMENT

Summary: The OpenStreetMap (OSM) community has over 1 million active contributors, around 50,000 of which are active each month. OSM is used in various scenarios and by Mega-Corporations as well as by humanitarian actors. Academics use the OSM data set in their research as supplementary data set, as well as study it as a primary research subject. Data quality and fitness-for-purpose have been a research focus of many scientists for more than a decade now and, with the increasing importance of OSM, are becoming even more relevant.

The Problem

- ▶ Most studies on OSM data quality are limited to small spatial extent and coarse temporal resolution → difficult to replicate or transfer the results to other regions
- ▶ Ideal: calculate data quality measures on a global context in a fine spatial and temporal resolution.
- ▶ Regular data updates → monitor improvement of OSM dataset

ohsome platform to help with different use cases

- flexible data analysis via the OSHDB
- ready-to-use web based ohsome API
- user interfaces such as ohsome HEX and OQT

OSM Dictionary: What is...

...Fitness for Purpose?

Quality of geodata often can not be assessed generally. It depends on the specific use case which of the many aspects of data quality is important (Completeness of data/attributes? Accurateness?)

...Intrinsic data quality measures?

Saturation measures can be used to determine the completeness of the map → the development of the data set over time can be evaluated

HeiGIT | HEIDELBERG INSTITUTE FOR GEOINFORMATION TECHNOLOGY

HeiGIT gGmbH
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg

CONTENT
Martin Raifer

DESIGN
Linda Sendlinger,
Dominik Neumann

The Solution: ohsome

- intrinsic methods for quality assessment
- in depth data analysis of spatio-temporal statistics of the OpenStreetMap data set
- Big Data technologies: apply methods up to global coverage
- Fine temporal resolution: possibility to investigate short time events (e.g. during disaster response)

The Results

- Quantitative statistics provide an overview of the status quo of OSM in arbitrary regions, and the temporal development of the data.
- Intrinsic data quality measures can be calculated easily..
- OSM data extraction allows to analyse OSM map features in GIS software and to use it for applications such as machine learning.
- In-depth analysis of the contributions to the OSM data provides insight into the map-making process, the users behind it and their interactions.

OSM Buildings 2020

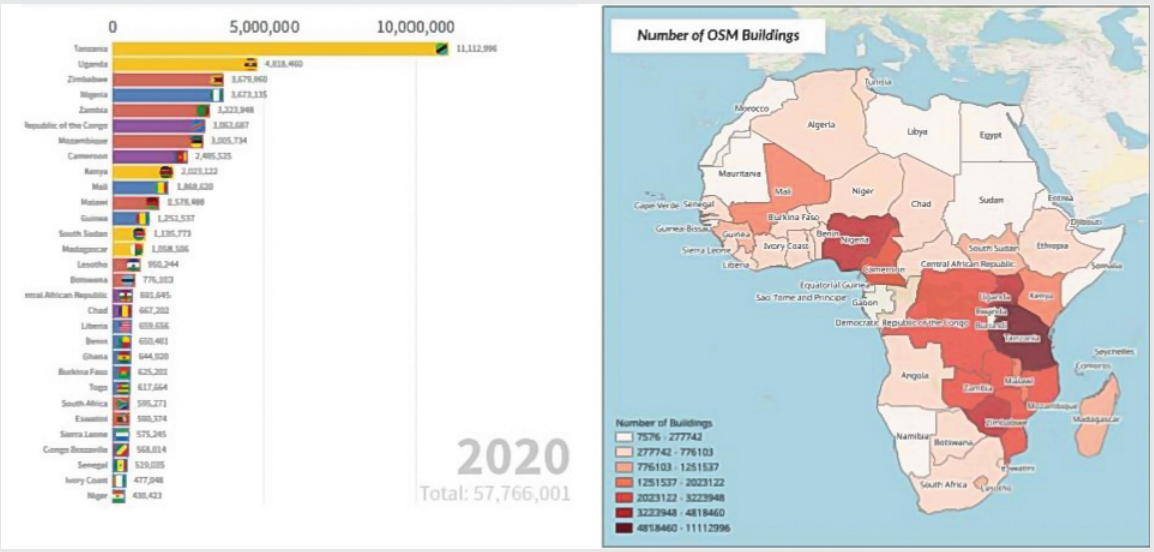

FIGURE: The figure illustrates the mapping of buildings in Africa in the year 2020

HeiGIT gGmbH
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg

CONTENT
Martin Raifer

DESIGN
Linda Sendlinger,
Dominik Neumann