5.4 Quality Assurance and Quality Control

Quality Assurance and Quality Control is the process by which mappers, and OpenStreetMap contributors in general, check data to ensure that all information uploaded to OSM meets high standards for usage and to prevent vandalism. As OpenStreetMap is a free and open platform that anyone can use and edit, it is critical to the sustainability of open data and OSM that everyone participates in the quality assurance and quality control process – from field data collection to data cleaning to long-term maintenance of existing OSM data.

The data quality dimensions identified through this framework as the highest priority for assessment of all datasets are:

- **Completeness**
- **Consistency**
- **Accuracy**
- **Timeliness**
- **Accessibility**

For each data quality phase, the following sets of questions should be answered:

Phase 1

Defining Needs & Requirements

Why are we collecting this data? Who will use it?

What data are we collecting?

What information will we and other users need to know about the data we collected?

What is the geographic scope of our data?

What is the temporal scope of our data collection? Does this timeline meet the needs of the users?

What quality metrics are required for the datasets? Is there an allowable margin of error or acceptable data gaps?

Have partners reviewed and agreed upon the needs and requirements of the project?

Does the scope of this project fall under the OSM Organized Editing Guidelines? If so, does it meet the requirements?

Phase 2

Defining Data Collection Methodology

Have permissions been secured to conduct data collection in the area of interest? Has proper community outreach been conducted with the community and stakeholders relevant for the area of interest?

Have team members been provided specialized and hands-on training for their role?

Has all software and hardware been selected based on the needs of the project? Has all software and hardware been tested?

Do data collection forms meet the need of the project and data model? Have data collection forms been trialed by surveyors and with survey subjects?

How will collected data be stored? What data protection strategies are in place?

Where will the data come from and what standard indicators will be used that can help achieve objectives?

Phase 3

Collecting Data & Producing Datasets

Are data collectors able to report issues and challenges in the field?

Is data quality being monitored during field data collection?

Is field data being safely secured on a daily basis?

Phase 4

Cleaning and Validating Datasets

Has raw data been reviewed for errors prior to upload or packaging? What tools were used in this process and why?

Do the cleaned data sets include private data or publicly identifiable information? If so, has the dataset been split into public and private sets?

Has data been properly uploaded to OSM with designated changeset comments and tags?

Has data uploaded to OSM been reviewed for additional quality assurance? What tools were used in this process and why?

Phase 5

Ensuring Sustainability & Providing Documentation

Has data been made available through an open and accessible program?

Has the availability of data been communicated to relevant stakeholders?

Have the workflows, tools, and processes used for the project been documented? Is this documentation open and available for users outside of HOT?

Have local stakeholders or other groups with potential for sustaining data use and data collection been trained in project methodology?

Quality Assurance Tools

The following are Quality Assurance tools commonly used in the HOT workflow. A detailed overview of these and other QA tools can be found at the Quality Assurance Tools Wiki. Different tools check for different errors and issues. If you want to:

- Check for attribute completeness, use MapCampaigner.
- Check for potential vandalism, use OSMCha.
- Check for contributions and upload issues by indivdual user, use OSMCha.
- Check for tagging issues, use Osmose AND JOSM Validation.
- Check for geometry issues, use Osmose AND JOSM Validation.

MapCampaigner

MapCampaigner is a tool developed by HOT to monitor attribute completeness for predefined areas of interest (AOI). Based on your data model, the tool checks and highlights any map features that are missing pre-defined tags within your AOI, and allows team validators to download and fix those features.

OSMCha

OSMCha, or the OpenStreetMap Changeset Analyzer, is a tool designed to review uploads and changes to OSM data, largely to prevent vandalism and bad edits made to map data. This tool allows users to filter by username, location, dates of upload, and other metadata features. OSMCha is useful for monitoring the progress of data cleaning and upload teams.

Osmose

Osmose is a tool that monitors multiple quality control issues in OSM. These include issues with feature geometry (such as overlapping buildings/nodes, incomplete features, and duplications), and also common tagging issues (such as missing, unsuitable or poorly formatted tags). More information about Osmose can be found at the Osmose OSM Wiki Page.

Note: to properly use this tool and view errors, you will need to use the Google Chrome internet browser. Firefox, Opera, Safari, and other browsers may not display the information correctly.

Steps to correct key issues identified by Osmose

1. Open JOSM on your computer.
2. Navigate to http://osmose.openstreetmap.fr in the Google Chrome internet browser.
3. Use the zoom, pan, and search features on the map to navigate to your area of interest.
4. Use the left panel to toggle common issues on and off.
5. Identified issues will appear in the map as pins matching the icons from the issues panel. Click on each pin to learn more about the object and associated issue(s).
6. In Osmose, locate the 'Export' button at the top of the page. Click 'Export', then 'JOSM'.
7. Fix the issues identified then re-upload to OSM.