Quality Assurance Concept

Learning Objectives:

- Understanding the definition of quality assurance
- Understanding the meaning of data quality assurance in OpenStreetMap

OpenStreetMap (OSM) is a free and open source map. OSM has a lot of data and information about various types of objects that exist on the surface of the earth. The data is collected and entered by OSM users spread all over the world. The variety of data entered and collected by OSM users often results in the quality of the data being different from one another and the lack of understanding of OSM users regarding data quality also affects the results of data available in OSM. Therefore, every *OpenStreetMap* user should not only add data, but also participate in maintaining data quality. In this discussion, you will learn what quality assurance is and the approach mechanism used to show how to maintain and improve the quality of OpenStreetMap data.

I. Definition of Data Quality Assurance

When you talk about data that comes to your mind it might be a series of related information about an event, object or phenomenon collected from a particular source. The more data available, the more need to checking the quality of the data before it can be processed. This aims to make the data to be processed free from *noise* or dirty data. Dirty data is data whose value shows an unnatural value compared to other data.

Data quality assurance is an activity to check the condition of the data in terms of accuracy, consistency, completeness, clarity, and anomalies. By paying attention to these aspects, a data can be said to be good if all aspects are met.

II. Quality Assurance Mechanisms

Since its launch in 2004 by *Steve Coast*, OSM has become an alternative and even the first choice for communities to complete spatial data in the areas they want. The ease, completeness, and flexibility of OSM allows the data available in OSM to be often reused for various purposes such as business, technology, social, even for humanitarian and disaster activities. This causes an increase in the amount of data available in OSM and will certainly affect the quality of the data itself. With its free and open nature, maintaining the quality of OSM data is very important so that later the use of OSM data for various needs especially those involving community interests such as humanity and disaster can be done in quality.

According to the book "*Assuring the quality of volunteered geographic information*" written by Goodchild and Li (2012), data quality assurance in participatory mapping activities can be done in broad terms with 3 mechanisms, namely:

1. Participatory Quality Assurance

Added by Surowiecki in his book entitled "*The Wisdom of Crowds*" (2004) the mechanisms of **data quality assurance** inmanner **a participatory** has several characteristics and advantages such as:

- multiple users can generate a collective agreement on a fault data found. This usually happens like misinformation or types of objects mapped in an area. Users who map the area can produce mutual agreement so that there are no more misinformation of the objects to be mapped.

- Some observations and experiences of an individual can strengthen the validity of observations and experiences of other individuals so as to reduce the risk of misunderstanding of a particular object.

- Together they can validate and check the quality of data and errors in certain areas so as to save time, energy and costs in conducting data quality assurance activities.

  This participatory quality assurance example has been carried out by the *Humanitarian Open-StreetMap Team* in collaboration with the *Resilience Network Initiative* (RNI) mapping the Purwodinatan Village, Central Semarang District, Semarang City. This activity involves elements of the community and is assisted by local students and surrounding committees in conducting mapping to inputting data mapped into *OpenStreetMap*.

2. Social Quality Assurance

Bit different but still related to quality assurance mechanisms in a participatory, **Social Quality Assurance** emphasizes the quality of the individual to do so. The more often a person validates and corrects the errors that exist in the data, then he will be more trusted to be able to guarantee data quality. Thus, someone who has a good reputation will be trusted to lead others to carry out quality assurance activities in an activity or project. In addition, social quality assurance can be carried out by forming a *working group* that becomes a media to notify and update activities related to data quality assurance. This can accelerate data quality assurance activities and make effective communication among members.

One example of this social quality assurance activity is that which has been carried out by the *OpenStreetMap Foundation* where they create several *working groups* that each group has a different discussion theme. For the quality of OpenStreetMap data discussed in the data discussion group (*Data Working Group*) where in the group discussed various matters related to the data in *OpenStreetMap* such as data licensing, data vandalism, disputes about data, as well as helping in determining usage policies for OpenStreetMap data. This data discussion group consists of members *OpenStreetMap foundation* who are contribute to *OpenStreetMap* and also some users who are recommended by members who are on the *OpenStreetMap foundation* itself. The activities from *Data Working Group* can be see by visit the wiki page for *Data Working Group*[1] and if you want to join to discuss, you can contact their members at data@osmfoundation.org.

3. Geographic Quality Assurance

The latest mechanism for **ensuring data quality** is by **geography**. This mechanism uses a geographic theory approach. Not all users can and may do quality assurance using geographic mechanisms. Only those who truly understand the geographic theory related to spatial data analysis such as *Spatial neighbors and auto-correlation (Moran Statistics)*, *Inferential Statistics and Analysis of Variance (ANOVA)*, and others. Therefore this mechanism is rarely used in participatory mapping activities, especially in *OpenStreetMap*. For example the documentation of data quality *OpenStreetMap* based on the geography approach in Indonesia was made by the Department of Geodesy and Geomatic Engineering, Faculty of Engineering, Gadjah Mada University in 2012, this documentation they sample data *OpenStreetMap* in several cities such as Jakarta, Surabaya, Bandung, Yogyakarta, Padang, and Dompu. These cities have been held mapping activities *OpenStreetMap* from 2011 to 2012 ranging from social mapping activities in the province of West Nusa Tenggara especially Dompu organized by ACCESS and the *Humanitarian OpenStreetMap Team* (HOT) to *Scenario Development for Contingency Planning*(SD4CP) conducted by *Humanitarian OpenStreetMap Team* (HOT) with *Australia Indonesia for Disaster Reduction* (AIFDR) and the National Disaster Management Agency (BNPB). Documentation on the quality of the data can be downloaded at: http://openstreetmap.id/docs/Final_Report-OSM_Evaluation_in_Indonesia_2012.pdf.

III. Data Quality Assurance in OpenStreetMap

The data quality assurance mechanism that has been described can be applied to the OpenStreetMap data quality assurance. Please note that quality assurance activities on *OpenStreetMap* are the responsibility of all users. Therefore, *OpenStreetMap* already has several technical guidelines in carrying out

---

[1]https://wiki.osmfoundation.org/wiki/Data_Working_Group

these activities using the three mechanisms that have been described. There are 2 types of tools used in conducting data quality assurance in *OpenStreetMap*, namely:

- Data Monitoring *OpenStreetMap*
- Tool for Error Detection in *OpenStreetMap*

Data monitoring tool is a tool that can be used to **view quality data *OpenStreetMap*** that you have created or data *OpenStreetMap* that is around the area of the mapping activity that you are doing. **The error detection tool** will help **show errors** contained in **data *OpenStreetMap*** so it will be easier to find data errors than having to search one by one manually. Some of these tools are made by users who develop *OpenStreetMap* according to their needs. You will also learn some errors indata *OpenStreetMap* and tools related to data quality assurance in *OpenStreetMap*.

SUMMARY

Now you must already understand what a data quality assurance, data quality assurance mechanisms, and ensuring the quality of data in *OpenStreetMap*. With the guarantee of data quality in *OpenStreetMap*, data quality in *OpenStreetMap* can be maintained and of course the results of data *OpenStreetMap* can be used by everyone.