JURNAL TEKNIK SIPIL DAN LINGKUNGAN| EISSN:2549-1407
https://jurnal.ipb.ac.id/index.php/jsil

Vol. 09 No. 02 Oktober 2024
DOI: 10.29244/jsil.9.2.221-230

How effective crowdsourced data during crisis emergency? A case study of the 2018 Palu-Donggala earthquake

Zainab Ramadhanis[1]* and Anjar Akrimullah[2]

[1]Department of Civil and Environmental Engineering, IPB University, Bogor, 16680, Indonesia
[2]Perkumpulan OpenStreetMap Indonesia, Jakarta, 12910, Indonesia
***Corresponding author: zainabr@apps.ipb.ac.id

Abstract: In disaster situations, updated geographic data is crucial for disaster relief efforts. OpenStreetMap (OSM) has demonstrated significant value in disaster response scenarios due to its capacity for rapid data collection and dissemination, since the 2010 Haiti earthquake. This study investigates the quality of OSM data during the 2018 Palu-Donggala earthquake, focusing on how contributor expertise affects data reliability and how effectively OSM data supports decision-making in emergencies. The research highlights the critical role of OSM in providing timely geospatial information, with 205 contributors mapping roads and buildings in Palu City and Donggala Regency within just three days of the earthquake. Our findings show that while road data had significant topological errors—7,085 errors primarily due to overshoots—building data had considerably fewer errors, with only 76 recorded. This disparity suggests that OSM data for buildings was of higher quality during the crisis. The preference of eight out of nine mapper types for building data over road data further underscores the value of OSM in emergencies, as experienced mappers tended to focus on features that were less error-prone. Moreover, the behavior of contributors was analyzed, showing that although a large number of contributors were inactive, most of the experienced contributors remained actively participating. This finding indicates the potential for inactive expert mappers to return and contribute in future crises. Additionally, the study assesses the rapid collection of data by OSM and its impact on decision-making. The National Disaster Management Agency of Indonesia (BNPB) and the ASEAN Coordinating Centre for Humanitarian Assistance (AHA Centre) effectively utilized the data to provide updates on fatalities, injuries, and displacement, facilitating a swift and equitable distribution of aid.

Keywords: 2018 Palu-Donggala earthquake, geographic data quality, OpenStreetMap

Submitted: September 14 2024
Revised : October 07 2024
Accepted : October 10 2024

1. Introduction

OpenStreetMap (OSM) has gained significant popularity as a Volunteered Geographic Information (VGI) platform over the years [1]. In August 2022, the total number of registered OSM accounts reached more than 9.9 millions. Moreover, the number of uploaded OSM data also increased, with more than 9.3 billion nodes, 1 million ways, and 15 million relations [2]. One of OSM's strengths lies in its large number of volunteers who bring valuable local knowledge to the platform [3] and potential to influence disaster management decision-making [4]. A pivotal moment in OSM's history was during the 2010 Haiti earthquake. On 12 January 2010, within 48 hours after the earthquake, more than 600 contributors mobilized to gather spatial data and construct a comprehensive basemap of Haiti through OSM. This effort proved crucial, as the Haitian government did not have an official basemap available at the time [5]. The rapidly collected data provided vital information for first responders, highlighting the importance of VGI in disaster response scenarios [6].

Since the Haiti earthquake, community participation played an instrumental role in crisis responses [7]; including OSM, which has become a key resource in several other crisis responses [8, 9], with the number of contributors increasing over the years. The OSM contributors participated in disaster mapping in pos-disaster through the Tasking Manager, a web tool developed by Humanitarian OpenStreetMap Team (HOT). This disaster mapping referred to remote mapping based on satellite imagery [10]; thus many people around the world can participate.

Another remarkable example occurred during the 2018 Palu-Donggala earthquake, which struck Central Sulawesi, Indonesia on 28 September. Within just three days, 205 volunteers had mapped key features such as buildings and roads in the affected regions of Palu City and Donggala Regency. Following the completion of this mapping effort, the Indonesian National Board for Disaster Management (BNPB) reported extensive damage, including 2,081 fatalities, 4,438 major injuries, 1,309 missing persons, 68,451 houses damaged, and 206,494 displaced [11]. The swift data collection enabled authorities to assess the situation more effectively, demonstrating OSM's capacity to provide rapid geospatial information in the aftermath of a disaster.

Despite the demonstrated success of OSM in crisis situations, some concern regarding the data quality as the different perspectives of an individual contributor may also influence the accuracy of data [12]. As a crowdsourced platform, OSM allows anyone to contribute, regardless of their formal training or educational background in mapping or geography. While the quick and global reach of OSM are invaluable in emergencies, the varying levels of expertise among contributors can lead to inconsistencies or inaccuracies in the data. The geospatial standards are crucial, especially during a crisis, as decisions based on inaccurate data could have severe consequences for disaster response and mitigation efforts.

Due to the limitation of references available for assessing OSM data, this study focuses on evaluating a key aspect of logical consistency, specifically topological consistency. Topological inconsistencies, such as misaligned roads or overlapping buildings, can be especially critical in disaster contexts, as they might impact response efforts. This research examines the accuracy of the topological features within the OSM data collected during the 2018 Palu-Donggala earthquake and explores contributor behavior by analyzing their mapping activities and preferences for specific features. Furtermore, the study aims to assess how well OSM data can support decision-making in emergency scenarios in data quality perspective.

2. Methodology

2.1. Research Location

The research was carried out between May and August 2024. Donggala Regency and Palu City, situated in Central Sulawesi, are located at coordinates 0.6829° S, 119.7324° E (Donggala Regency) and 0.8989° S, 119.8500° E (Palu City). The map of the research locations is shown in **Figure 1**.

2.2. Tools and Data

This study employed QGIS 3.28, C++ on Terminal (Mac), the "How did you contribute to OSM" (HDYC) website tools, and Microsoft Excel as its primary tools. QGIS was used to evaluate the quality of geographic data according to ISO 19157:2013 standards. C++ on Terminal (Mac) facilitated the extraction of changesets from OpenStreetMap data to analyze contributor behavior. The HDYC tool helped identify mapper types, feature preferences, and activity status. Microsoft Excel was utilized for statistical analysis of contributor behavior based on their mapping activities in OpenStreetMap. Additionally, the study used geographic data and changeset comments from OpenStreetMap for the Palu and Donggala regions, which were sourced from Geofabrik (https://download.geofabrik.de/asia/indonesia.html) and OSM Planet (https://planet.openstreetmap.org/planet/2024/).

Figure 1. Research Location

2.3. Experiments

2.3.1 Geographical Data Quality Evaluation: Logical Consistency

The International Organization for Standardization (ISO) defines ISO 19157:2013 as a standard for geographical data that takes into account users' needs, distinguishing it from earlier versions like ISO/TS 19138:2006, ISO 19114:2003, and ISO 19113:2002 [13]. ISO 19157:2013 addresses data quality by outlining key concepts, establishing quality conformance levels in data specifications or user requirements, specifying quality aspects, and providing methods for evaluating and reporting data quality.

It identifies six data quality elements: completeness, positional accuracy, logical consistency, thematic accuracy, and temporal quality as internal elements assessed based on data produced by the data producer, and usability as an external element evaluated according to user needs. The standardization describes data quality evaluation by comparing datasets to a universe of discourse as the reference, focusing on elements like completeness, positional accuracy, thematic accuracy, and temporal quality. Logical consistency, however, is evaluated through spatial relationships within a geographical region rather than direct comparison [14]. Logical consistency consists four data quality elements: conceptual consistency, domain consistency, format consistency and topological consistency [13]. Given the limitations of references, only topological consistency was evaluated.

Topological consistency was evaluated to reveal potential sources of data inconsistencies, such as missing nodes, pseudo nodes, undershoots, overshoots, duplicate lines, and missing or duplicate reference points, as shown in **Figure 2**. To identify topological inconsistencies, we employed the Topology Checker tool available in the QGIS plugin, which needs to be installed beforehand. For road data, we applied five topology rules for polylines: no dangles, no duplicates, no invalid geometries, no multi-part geometries, and no pseudos. For building data, we similarly applied five topology rules for polygons: no duplicates, no gaps, no invalid geometries, no multi-part geometries, and no overlaps. A lower number of errors indicates higher data quality.

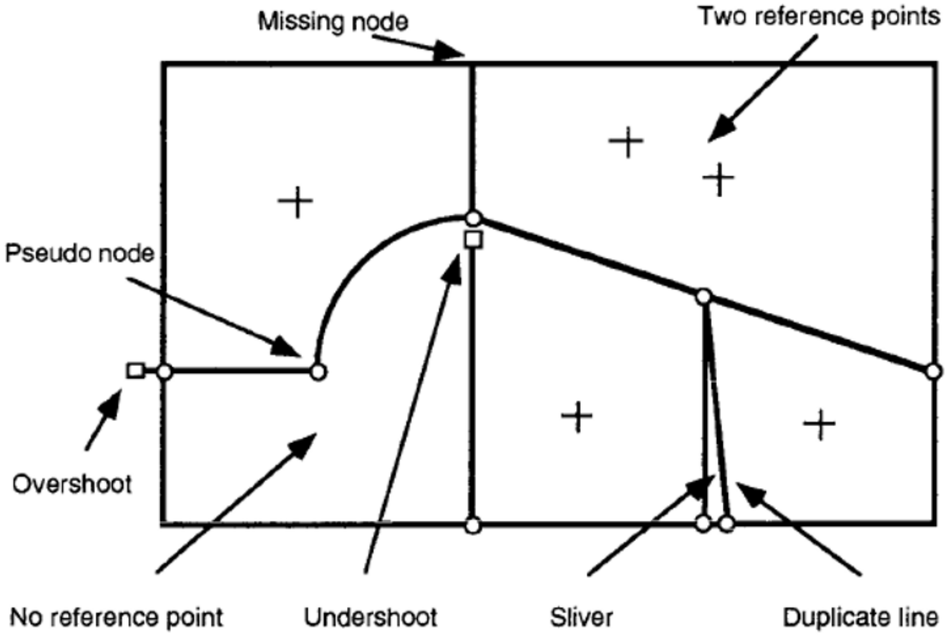

Figure 2. Conditions of error in digital spatial data sets [15]

2.3.2 OpenStreetMap Contributors Behavior

To analyze the behavior of volunteers who contributed to OSM after the 2018 Palu-Donggala earthquake, we adopted the work from Van Exel et al. [16] and Bégin et al. [17], who highlighted that OSM contributors are diverse, with varying quality, motivations, and preferences. This diversity means that evaluating spatial data quality alone is insufficient for crowdsourced data. They introduced the concept of Crowd Quality, which quantifies crowdsourced geospatial information in two dimensions: user quality and feature quality. User quality encompasses the local knowledge, experience, and recognition of OSM users, while feature quality refers to the spatial quality aspects like lineage, consistency, and topology.

In this study, we retrieved the OpenStreetMap changeset history from September to October 2018 for Palu City and Donggala Regency. This data includes the date, ID, coordinates of updates, usernames of contributors, and comments describing each change. Since this data is compressed in bz2 format, it is not directly readable. Furthermore, because it covers a broad geographic area, we needed to extract only the relevant data for our study area and convert it into XML format for easier interpretation. We used the Open-Source C++ package libosmium for this extraction process on macOS Sonoma 14.4.1.

We then analyzed each contributor's behavior by using their usernames on the "How did you contribute to OSM?" (HDYC) web page, which offers detailed information about OSM contributors. Founded by Pascal Neis in 2010, HDYC provides statistical data on contributors based on OSM data dumps. In our analysis, we manually entered usernames into the HDYC page to gather data on user quality and feature preferences. User quality was evaluated based on contributor status (active or inactive) and mapper type, while feature preferences were determined by examining the frequency of features associated with nodes, ways, and relations.

3. Results and Discussions

Logical consistency, a key aspect of ISO 19157 quality standards, pertains to the product's quality from the producer's perspective [13]. Given that OpenStreetMap depends on its contributors, understanding their behavior is crucial for evaluating the overall quality of the data.

3.1. Logical Consistency

Logical consistency is evaluated to describe the organization of the data structure in digital spatial data [15], which includes features, relationships, and attributes. ISO 19157 defines three elements of logical consistency; however, this research focused solely on assessing the topological consistency of road and building datasets. The study employed the Topology Checker plugin in QGIS. The results of the topological error analysis for road data is presented in **Table 1**. The analysis shows that 81.2% of the topological errors in road data are due to dangling ends, indicating that 5,753 road features are overshot in Palu City and Donggala Regency.

Many contributors may have missed these overshoot errors (dangling ends) because they used inconsistent data sources, some used Bing Imagery, some used Maxar Imagery, whose precision and detail varied. Thus, road data is not always matched with satellite imagery. The quality of OSM data also depends on the tools and experience of contributors, as formal training is not mandatory; in the absence of quality control, errors like overshoots may continue to occur. Additionally, quick data collecting is one of OSM's main advantages. However, if the process does not handle inconsistencies adequately, discrepancies may appear when new data is added or updates are combined, which could cause overshoots if the new data does not match the old data.

Tabel 1. Logical consistency of road data

Error type	Total
Dangling end	5,753
Duplicate lines	0
Invalid geometries	0
Multi-part geometries	11
Pseudo nodes	1,321

Table 2 also highlights the logical consistency of building data. We applied five rules on polygon layers with Topology Checker, must not have gaps, must not have duplicates, must not have invalid geometries, must not have multi-part geometries, and must not overlap. Compared to road data, building data exhibited fewer errors, with a total of 76 errors, and 60.5% of the buildings were found to be overlapping. Overlapping buildings are generally easier to identify than overshot roads. Additionally, the analysis revealed that 15.4% of contributors used the iD editor for mapping, while 79.1% preferred JOSM. Unlike JOSM, the iD editor does not provide a warning for topology errors before the data is uploaded to the OSM database.

Tabel 2. Logical consistency of road data

Error type	Total
Duplicate polygons	20
Invalid geometries	9
Multi-part geometries	1
Overlapping	46

3.2. The Pattern of OpenStreetMap Contributors Behavior

This research assumes that each OSM account represents a single contributor. We identified 205 contributors who were involved in mapping the earthquake-affected areas of the 2018 Palu-Donggala disaster, though six contributors have since deleted their accounts. **Table 3** presents the quality of contributions, categorized as active or inactive. Activity status is determined by whether a contributor has participated in mapping more than four times within a 12-month period. While 68.8% of contributors were inactive, and only 28.23% remained active, it is noteworthy that 61.7% of the inactive group were experienced contributors with over 100 changesets. This suggests that inactive expert mappers could potentially return to contribute, especially in the aftermath of a major disaster, provided the information is widely published, as was the case here.

Tabel 3. User quality of OpenStreetMap's contributors during post-disaster

Type Mapper	Active Contributors	Inactive Contributors
Hit-and-Run	0	39
Newbie	0	15
Casual mapper	10	32
Great mapper	10	15
Heavy mapper	17	25
Super mapper	2	0
Legendary mapper	16	15
Fantastic mapper	2	0
Mega mapper	1	0
Total	**58**	**141**

Table 4 illustrates the preferences for different feature types, showcasing the top three features contributed by various mappers. Our results indicate that eight out of nine mapper types prioritized mapping buildings during this effort. This emphasis on building mapping was primarily motivated by the urgent need from the National Disaster Management Agency to quickly assess the number of damaged houses and displaced individuals for public reporting.

Tabel 4. The top three object preferences for each type of mapper

Type Mapper	Top 3 Feature Type Preference OSM Contributors		
	First	**Second**	**Third**
Hit-and-Run	Building	Name	Name
Newbie	Building	Highway	Highway
Casual mapper	Building	Highway	Highway
Great mapper	Building	Highway	Name
Heavy mapper	Building	Highway	Address
Super mapper	Building	Address	Highway
Legendary mapper	Building	Address	Highway
Fantastic mapper	Building	Building	Address
Mega mapper	Highway	Highway	Power

Another significant observation concerning the mapping duration of volunteers is the typical time taken for contributions, as depicted in **Figure 3**. The basic descriptive statistics in this figure illustrate the mapping duration (Y-axis) against the time for each uploaded changeset (X-axis). The longest recorded time for editing and uploading to the OSM server was approximately 1,415 minutes, or nearly 24 hours, while the shortest duration was less than 1 minute, with an average contribution time of about 19 minutes per edit.

Further analysis indicates that changesets lasting over 1,000 minutes represent only 115 out of a total of 70,524 changesets, or about 0.001%. This occurrence is not uncommon, as contributors occasionally forget to upload their changes for extended periods. In comparison, there were 1,624 changesets (approximately 0.02%) where editing sessions lasted more than 3 hours but less than 1,000 minutes. Additionally, changesets with editing durations between 1 and 3 hours made up 4,821, or roughly 0.06%. Changes that took longer than 30 minutes accounted for 9,219 changesets, about 0.13%.

Moreover, Figure 3 shows that nearly 99.9% of all changesets exhibit a consistent mapping duration. This implies that contributors are committed to completing their edits efficiently and responsibly, maximizing their time for humanitarian mapping initiatives. The average duration of 19 minutes per changeset supports this observation. It is also noteworthy that each OSM grid covers relatively small areas with low object density, especially in rural regions.

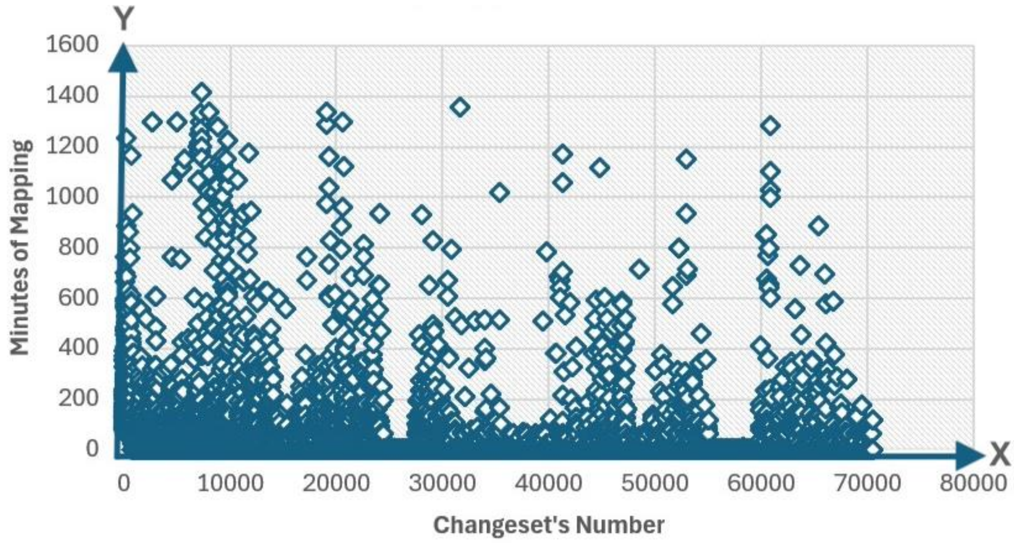

Figure 3. Mapping Duration Scatter Graph

From this descriptive analysis, we can conclude that volunteers mapping for humanitarian purposes on OSM show high dedication, carefully managing their time to contribute data in disaster-affected areas. Their time commitment is neither excessively long nor short, yet they remain focused on completing their tasks, with an average session lasting 19 minutes.

3.2. The Usability of OpenStreetMap Data for Post-Disaster in 2018 Palu-Donggala Earthquake

The use of OpenStreetMap (OSM) data following the 2018 Palu-Donggala earthquake demonstrated its critical role in post-disaster recovery, especially in terms of quickly gathering and distributing information. Within three days, 205 global contributors mapped essential features such as buildings and roads in Palu City, Donggala Regency, and surrounding regions. This rapid mapping helped authorities and humanitarian agencies assess the infrastructure damage. Since OSM data is openly accessible, government agencies, NGOs, and individuals could use it to meet their specific needs. Continuous map updates also proved crucial, allowing responders to adapt to the dynamic conditions of disaster environments.

The spatial data collected via OSM supported agencies like Indonesia's National Disaster Management Agency (BNPB) and the ASEAN Coordinating Centre for Humanitarian Assistance (AHA Centre) in evaluating affected areas, particularly those hit by liquefaction. This data enabled more informed decisions, helping coordinate relief and recovery efforts efficiently. By having accurate and up-to-date maps helped humanitarian organizations plan routes, manage logistics, and deliver aid like shelter, food, and medical supplies to the hardest-hit areas.

The collaborative nature of OSM encouraged global engagement, with volunteers worldwide mapping affected regions based on satellite imagery as news about 2018 Palu-Donggala earthquake spread globally through media such as BBC, CNN and Al Jazeera. This widespread attention enabled OSM to rapidly scale its disaster response efforts, accelerating the mapping process during the crisis.

For example, as shown in **Figure 4**, areas like Balaroa, Petobo, and Biromaru were significantly affected by liquefaction. With mapping data from OSM, BNPB calculated that around 430.7 hectares of land were inundated, affecting 3,773 buildings. This detailed mapping also helped the Ministry of Health assess the need for healthcare workers in the region, including specialists such as physicians, pediatricians, anesthesiologists, and more. This example illustrates the critical impact of OSM in facilitating disaster response, both locally and internationally.

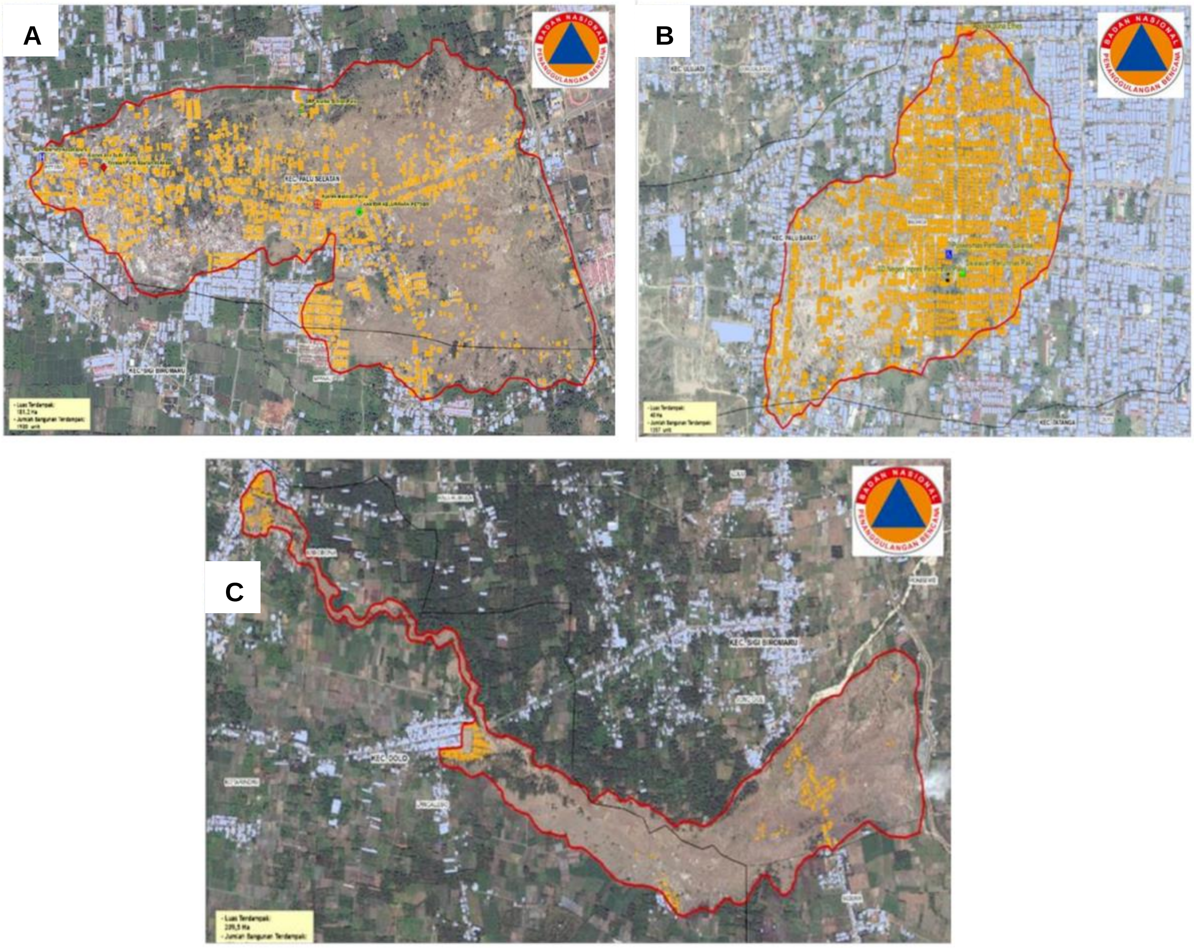

Figure 4. Petobo area (A); Balaroa area (B); Biromaru area (C)

4. Conclusion

This study addresses the critical issue of OSM data quality during crisis emergencies, emphasizing the need for rapid data collection. Our findings reveal that 205 contributors managed to map roads and buildings within three days, despite an average mapping time of just 19 minutes. Although road data contained 7,085 topology errors—primarily due to overshots—errors in building data were significantly fewer, with only 76 errors recorded. This suggests that OSM data, particularly for buildings, proved to be a valuable resource during the crisis. The preference of eight out of nine mapper types for building over road mapping further supports the effectiveness of OSM data in crisis situations. This preference might be linked to the fact that building data had fewer errors, indicating that experienced mappers could contribute more effectively to certain types of data. However, further investigation is needed to fully understand how different levels of expertise affect data quality across various feature types.

Moreover, the effectiveness of OSM data in supporting decision-making during the 2018 Palu-Donggala earthquake was evident. The rapid data collection enabled the Disaster Management Agency and the AHA Centre to promptly update the public on fatalities, injuries, missing persons, damaged houses, and displaced individuals. This timely information was crucial for both BNPB and the AHA Centre to coordinate a swift and equitable response, ensuring that aid was distributed efficiently to the most affected communities. To improve the effectiveness of OSM data in future emergency responses, we recommend implementing more rigorous data validation processes to minimize errors, especially in critical areas such as roads. Additionally, future studies should investigate the impact of mapping duration on data quality, particularly for road data.

Acknowledgement
The authors extend their gratitude to the OpenStreetMap contributors, with special thanks to the 205 individuals who participated in mapping efforts during the 2018 Palu-Donggala earthquake.

Daftar Pustaka

[1] Neis, P., & Zielstra, D. (2014). Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future internet*, *6*(1), 76-106.
[2] OpenStreetMap. (2024, August 15). Stats. https://planet.openstreetmap.org/statistics/data_stats.html
[3] Birkmann, J., & von Teichman, K. (2010). Integrating disaster risk reduction and climate change adaptation: key challenges—scales, knowledge, and norms. Sustainability Science, 5(2), 171-184.
[4] de Albuquerque, J. P., Eckle, M., Herfort, B., & Zipf, A. (2016). Crowdsourcing geographic information for disaster management and improving urban resilience: an overview of recent developments and lessons learned. European handbook of crowdsourced geographic information, 309-321.
[5] Radford, T. (2020). Haiti 10 Years Later: Growth of a Humanitarian Mapping Community. Humanitarian OpenStreetMap Team. Retrieved 19 August 2024 from https://www.hotosm.org/updates/haiti-10-years-later-growth-of-a-crisis- mapping-community/.
[6] Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. International Journal of Digital Earth, 3(3), 231-241.
[7] Soden, R. (2014). 4 Years On, Looking Back at OpenStreetMap Response to the Haiti Earthquake. Retrieved 19 August 2024 from https://blogs.worldbank.org/en/latinamerica/4-years-looking-back-openstreetmap-response-haiti-earthquake
[8] Neis, P., Singler, P. & Zipf, A. (2010). Collaborative mapping and emergency routing for disaster logistics-case studies from the Haiti earthquake and the UN portal for Afrika. Geoinform. Forum Salzburg 4, 239–248.
[9] Aitsi-Selmi, A., Murray, V., Wannous, C., Dickinson, C., Johnston, D., Kawasaki, A., ... & Yeung, T. (2016). Reflections on a science and technology agenda for 21st century disaster risk reduction: Based on the scientific content of the 2016 UNISDR science and technology conference on the implementation of the Sendai framework for disaster risk reduction 2015–2030. International Journal of Disaster Risk Science, 7, 1-29.
[10] Albuquerque, J. P. D., Herfort, B., & Eckle, M. (2016). The tasks of the crowd: A typology of tasks in geographic information crowdsourcing and a case study in humanitarian mapping. Remote Sensing, 8(10), 859.
[11] AHA Centre - Emergency Operations Centre (EOC) in cooperation with National Disaster Management Authority (BNPB; Data & Information Centre, National Assisting Post). (2018). SITUATION UPDATE NO. 15 - FINAL M 7.4 EARTHQUAKE & TSUNAMI Sulawesi, Indonesia. Retrieved 19 August from https://ahacentre.org/wp-content/uploads/2018/10/AHA-Situation_Update-no15-Sulawesi-EQ-rev.pdf.
[12] Ather, A. (2009). A Quality Analysis of OpenStreetMap Data, M. Eng Dissertation, Department of Civil, Environmental & Geomatic Engineering ...].
[13] International Organization for Standardization. (2013). ISO 19157 Geographic information: data quality. In. Geneva, Switzerland: ISO.

[14] Theobald, D. M. (2001). Topology revisited: representing spatial relations. *International Journal of Geographical Information Science*, *15*(8), 689-705.
[15] Kainz, W. (1995). Chapter six: Logical consistency in Elements of spaital data quality by Guptill and Morrison. In. Elsevier.
[16] Van Exel, M., Dias, E., & Fruijtier, S. (2010). The impact of crowdsourcing on spatial data quality indicators. Proceedings of the GIScience 2010 Doctoral Colloquium, Zurich, Switzerland.
[17] Bégin, D., Devillers, R., & Roche, S. (2013). ASSESSING VOLUNTEERED GEOGRAPHIC INFORMATION (VGI) QUALITY BASED ON CONTRIBUTORS' MAPPING BEHAVIOURS. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-2/W1, 149-154. https://doi.org/10.5194/isprsarchives-XL-2-W1-149-2013.