Towards assessing the quality of volunteered geographic information from OpenStreetMap for identifying critical infrastructures

Benjamin Herfort
GIScience Chair, Heidelberg University, Germany
Herfort@stud.uni-heidelberg.de

Melanie Eckle
GIScience Chair, Heidelberg University, Germany
Eckle@stud.uni-heidelberg.de

João Porto de Albuquerque
GIScience Chair, Heidelberg University, Germany
Dept. of Computer Systems/ICMC, University of Sao Paulo, Brazil,
jporto@icmc.usp.br

Alexander Zipf
GIScience Chair, Heidelberg University, Germany

ABSTRACT

Identifying the assets of a community that are part of its Critical Infrastructure (CI) is a crucial task in emergency planning. However, this task can prove very challenging due to the costs involved in defining the methodology and gathering the necessary data. Volunteered Geographic Information from collaborative maps such as OpenStreetMap (OSM) may be able to make a contribution in this context, since it contains valuable local knowledge. However, research is still due to assess the quality of OSM for the particular purpose of identifying critical assets. To fill this gap, this paper proposes a catalogue of critical asset types, based on the analysis of different reference frameworks. We thus analyze how good the emergent OSM data model is for representing these asset types, by verifying whether they can be mapped to tags used by the OSM community. Results show that critical asset types of all selected sectors and branches are well represented in OSM.

Keywords

Emergency planning, Critical Infrastructure, OpenStreetMap, Volunteered Geographic Information, Disaster Management

INTRODUCTION AND BACKGROUND

With the increasing frequency of disasters caused by natural hazards, emergency planning becomes a crucial task for every municipality. For instance, by the year 2015 the development of flood risk maps as well as alarm and emergency plans will be compulsory within the European Union as decided in the Flood Directive 2007/60/EG (Parliament 2007).

In this context, the concept of Critical Infrastructure (CI) is commonly used for referring to objects that must be considered in planning since they have a critical role for society, either because of their importance for the functioning of a society or due to their significance for emergency management in the case of a disaster (Bouchon 2006; Organization of American States 1991). The main idea behind this concept is to focus on essential assets in emergency planning, which could cause important damage in case of different types and severities of hazards. However, there is no consensual definition of which infrastructures are critical (Haemmerli & Renda 2010, Comes, Bertsch & French 2013), let alone consensual methodologies to identify the individual assets that are part of a critical infrastructure (Moteff & Parformak 2004). Since particular regions have idiosyncratic conditions to be considered, it is usually expected that each municipality will identify their own critical assets for developing their emergency plans (Organization of American States 1991). Nevertheless, this task can prove very challenging to municipalities, since they may lack resources for defining customized methodologies for asset identification, as well as for building a comprehensive and up-to-date information basis.

Collaborative maps such as OpenStreetMap (OSM) were suggested as a potential source of Volunteered Geographic Information (VGI) for identifying elements at risk of a community (Schelhorn et al. 2014). OSM is collaboratively defined by citizens according to the Wikipedia principle, thus it may contain valuable local knowledge and thereby present an alternative for the costly methods of obtaining and maintaining official data. There are already reported cases in which OSM data proved to be of major use for disaster risk management (Neis et al. 2010; Soden et al. 2014). This is in line with the recent trend of citizens participating in disaster response by creating own applications to gather and exchange knowledge or add their information to official maps that are opened up for the public (Turoff et al. 2013). Nevertheless, the quality and credibility of information produced by volunteers is still a major concern for emergency agencies and disaster management professionals.

A growing body of research has been conducted to analyze the quality of OSM data (Arsanjani et al. 2013; Barron et al. 2013; Haklay 2010; Hecht et al. 2013; Neis & Zipf, 2012). While these research studies showed that the quality of OSM data is in many regions comparable to official or commercial data sets (Neis & Zipf, 2012), it is also clear that the OSM data poses challenges regarding the heterogeneity and inconsistency of its semantic data structure (Mooney & Corcoran 2012). OSM does not have a fixed semantic model for objects such as an ontology or catalogue. Instead, OSM elements are assigned “tags” that are represented in a freely-chosen key-value-structure, also called “features”.

Although there is no compulsory feature catalogue, there are general guidelines regarding the attributes or “tags” to be used for specific object types. These tagging guidelines are constructed by the OSM community and can vary for different regions in the world and specific applications, e.g. the list of Map Features Germany[1]. Furthermore, new tags have been proposed by the OSM community for adding information on a wide range of topics related to disasters, which are consolidated in a Humanitarian Data Model[2] (Neis et al. 2010). However, there is currently a lack of studies evaluating to what extend this information is useful for effectively meeting the requirements of specific applications and decision makers, and particularly for the domain of emergency planning.

To fill this gap, the objective of this paper is to present first results towards a method for assessing the quality of OpenStreetMap for the specific purpose of identifying assets of critical infrastructure in support of emergency planning.

The remainder of the paper is organized as follows. First, the approach and method employed in this study are described. Afterwards, first results of the analysis of the OSM data structure are presented. Finally, the paper concludes with final remarks and points out directions for future work.

[1] http://www.wiki.openstreetmap.org/wiki/DE:Map_Features
[2]http://wiki.openstreetmap.org/wiki/Humanitarian_OSM_Tags/HDM_preset
Acess: January 30, 2015

APPROACH AND METHOD

This paper addresses the following research question:

RQ: How good is the OSM data structure for identifying Critical Infrastructure assets?

In order to answer this question, our approach starts with the examination of existing frameworks about Critical Infrastructure for deriving a catalogue of asset types of critical infrastructures that should be considered in emergency planning. This catalogue is then used to verify whether the data structure of OpenStreetMap is capable of representing the asset types contained in the catalogue. In doing so, we adopt a top-down approach, which starts from the information needs of the domain of emergency planning and go down to evaluate whether these needs are fulfilled by the volunteered geographic information of OpenStreetMap.

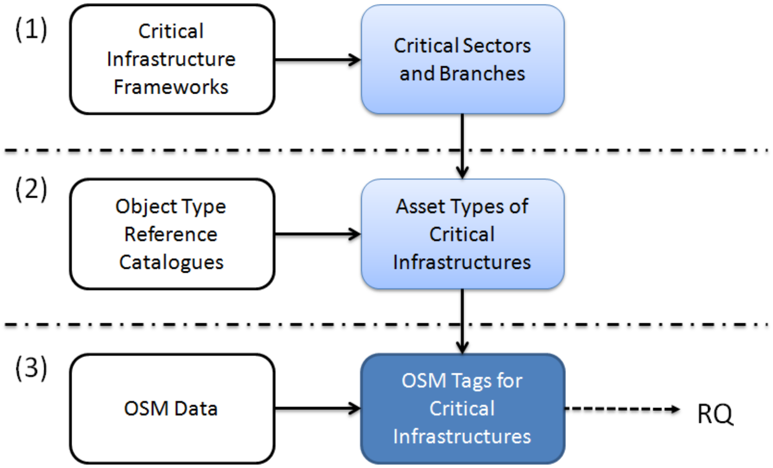

Figure 1. Research Approach

Figure 1 schematically depicts our methodology which is divided into three main components: (1) definition of a catalogue of critical sectors and branches using existing frameworks; (2) definition of a catalogue of asset types contained in the critical sectors and branches, based on object type reference catalogues; (3) definition of a catalogue of OSM tags used for assets of Critical Infrastructure. Finally, based on the catalogue of OSM tags, we answer our research question by verifying to what extend the OSM data structure (i.e. the existing tags) is capable of representing assets of critical infrastructures.

(1) Definition of a catalogue of critical sectors and branches

The German *national strategy for protection of critical infrastructures* (Bundesministerium des Innern) functioned as a basis for the definition of the catalogue of critical sectors and branches. The acquired sectors and branches were compared to sectors and branches in international frameworks dealing with critical infrastructures. These frameworks included the *critical infrastructure resilience strategy* of the Australian government (2010), the *council directive on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection* of the European Union (2009) and the *national strategy for the physical protection of critical infrastructures and key assets* of the US government (Bush 2003). Although there were differences in the distinction between sectors and branches, especially comparing US and European frameworks, we matched all initial sectors and branches to corresponding categories in the reviewed frameworks. Doing so, we generated a verified catalogue of critical sectors and branches.

(2) Definition of a catalogue of asset types of critical infrastructures

In a second step we merged object type reference catalogues and the previously developed catalogue of critical sectors and branches to derive a catalogue of asset types of critical infrastructures. To obtain a detailed selection of asset types we regarded universal and disaster management specific reference catalogues. They are described in the following.

The ALKIS reference catalogue includes all object types that are registered in the German cadastre. This catalogue is used by municipalities as well as by corporations on a national level. The LUBW framework was developed by the Department for Environment, Climate and Energy Baden-Wuerttemberg following the Flood Directive 2007/60/EG (Parliament 2007). The reference catalogue focuses on flooding events and contains object types for risk assessment and emergency planning. Since the LUBW reference refers to asset types in the ALKIS catalogue for the vast majority of asset types, we synthesize the catalogues and name then "ALKIS" in the following of this paper.

The HAZUS reference catalogue was developed by the Federal Emergency Management Agency of the United States Department of Homeland Security for multi hazard loss estimation. It includes information on General Building Stock, Essential Facilities, High Potential Loss Facilities, User Defined Facilities, Transportation Systems and Utility Systems.

To define asset types of critical infrastructures we selected asset types from the reference catalogues according to the following criterion. Do the asset types in the reference catalogue refer to any of the sectors and branches ascertained in the previous step? We independently classified all asset types within the ALKIS and HAZUS references according to this criterion. One should empathize that we did not classify the asset types according to their actual criticality. We rather generated a catalogue of potentially critical asset types.

Next, identical object types from different reference catalogues were merged. The final catalogue contains unique asset types of critical infrastructures that are listed in at least one object type reference catalogue.

(3) Definition of a catalogue of OSM tags

In the third step we added OSM tags to the asset types of critical infrastructures obtained in the second step. For doing so, we initially searched within the OSM Map Features, the OSM Map Features Germany and Humanitarian Data Model tagging guidelines to identify the proposed key-value-structure.

The OSM Map Features are a tagging guideline established by the OpenStreetMap community. The OSM Map Features Germany is customized version of the above. Both function as an informal standard, the former internationally, the latter especially for describing features in Germany. The Humanitarian Data Model is a tagging guideline developed for emergency management purposes, nevertheless it is not used as widely as the other two guidelines.

Regarding asset types that are not listed within these tagging guidelines, we further included other tags that are frequently used by the OSM community. We used the taginfo api[3] to determine whether the proposed key-value-structure is adopted more than 1000 times by the OSM contributors and added this feature to the catalogue if true. We applied this due to the fact that new features in OSM can be introduced without proposing them on a wiki page, but solely by using them. For this reason there are features that are used widely but not listed in the Map Features tagging guidelines.

FIRST RESULTS

The results of step one and two of the method described in the previous section are presented in Table 1. The resulting catalogue of asset types of critical infrastructures consists of 9 sectors comprising 27 branches. Within the branches there are 342 asset types. "Transport" contains by far the most asset types (108), while "Finance and Insurance Industry" and "Telecommunication and Communication Technology" are less represented (5, 12). The number of asset types within the other sectors is less extreme (21-56).

Table 2 shows the number of asset types contained in the different reference catalogues. While the reference catalogue from HAZUS only contains less than half of all asset types (160), ALKIS and OSM include considerably more asset types (267, 239).

[3] https://taginfo.openstreetmap.org/ , Access: January 30, 2015

Sectors	**Branches**
Energy (28)	Electricity (13); Oil (8); Gas (7) *e.g. power plant, substation, oil storage tank, gas pipeline*
Telecommunication and Communication Technology (12)	Telecommunication (5); Communication Technology (7) *e.g. communication tower, antenna, telecommunication provider*
Transport (108)	Aviation (14); Maritime Navigation (1); Inland Water Navigation (35); Railway (25); Road (29); Logistics (4) *e.g. airport, harbor, station, railway bridge, railway tunnel, motorway, highway*
Health (21)	Medical Care (18), Medicine and Vaccine (2); Laboratory (1) *e.g. hospital, social facility, pharmacy, medical laboratory*
Water (29)	Water Supply (21); Sewage Water Disposal (8) *e.g. water works, water reservoir, hydrant, wastewater plant*
Nutrition (56)	Food Industry (33); Food Trade (23) *e.g. supermarket, department store, restaurant, farmland, flour production*
Finance and Insurance Industry (5)	Banks (4); Insurance (1) *e.g. credit institution, bank, atm, money exchange, insurance company,*
State and Administration (47)	Government and Administration (28); Parliament (1); Justice (4); Emergency (14) *e.g. Parliament, townhall, primary school, prison, police, fire station*
Media and Culture (36)	Broadcasting (2); Cultural Assets (12); Symbolic Monuments (22) *e.g. radio station, tv station, theatre, opera, church*

Table 1. Catalogue of asset types of critical infrastructures, number of asset in each branch in brackets

# OSM	# HAZUS	#ALKIS	# Catalogue
239	160	267	342

Table 2. Number of asset types contained in reference catalogues

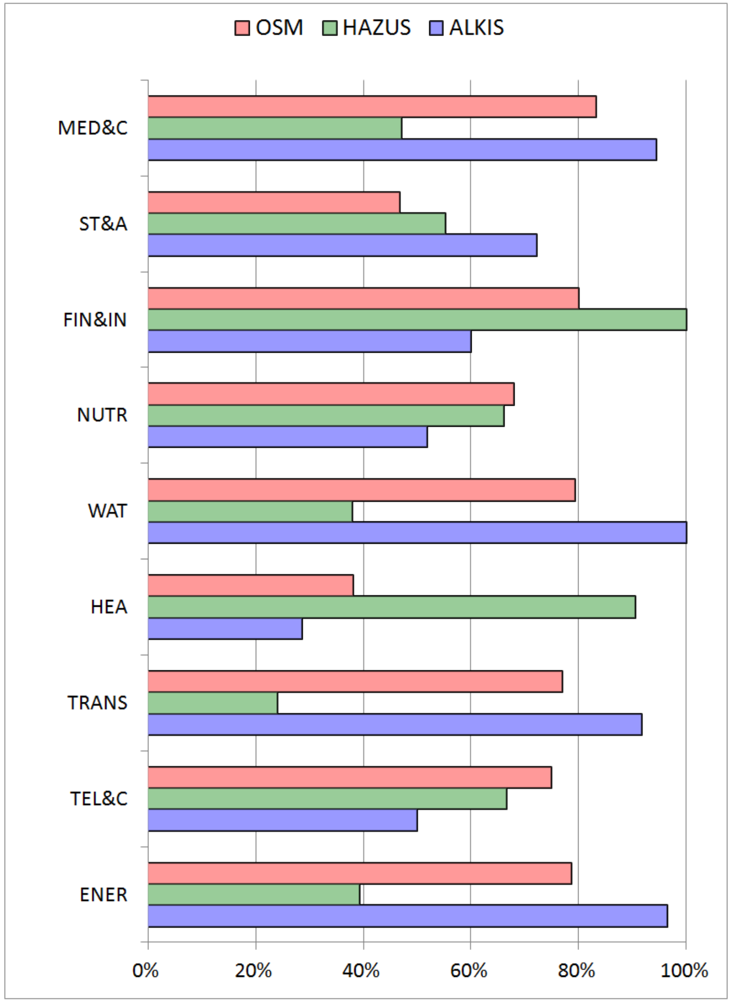

Figure 2. Share of asset types for different reference catalogues and critical Sectors in percent

Figure 2 provides a more detailed view on these differences between the reviewed reference catalogues. While HAZUS is by far the most suitable source for the sectors “Health” and “Finance and Insurance Industry”, ALKIS performs best in the sectors “Media and Cultures”, “Water”, “Energy” and “Transport”. In the sectors “Water”, “Transport” and “Energy” HAZUS can provide only less than half the total identified asset types. OSM provides more than 75% of all asset types for the sectors “Media and Culture”, “Finance and Insurance Industry”, “Water” and “Transport”. However, OSM and ALKIS cover only less than 50% of all asset types of the sector “Health”.

Furthermore, the results from OSM are more homogeneous considering different sectors (min: 38.1%, max: 83.3%) than the results from ALKIS (min: 28.6%, max: 100%) and HAZUS (min: 24.1%, max: 100%).

Finally, we examined the relationship among our catalogue, OSM, HAZUS and ALKIS using Venn diagram (See Figure 3). The bold dashed circles cover all asset types of critical infrastructures listed in our catalogue.

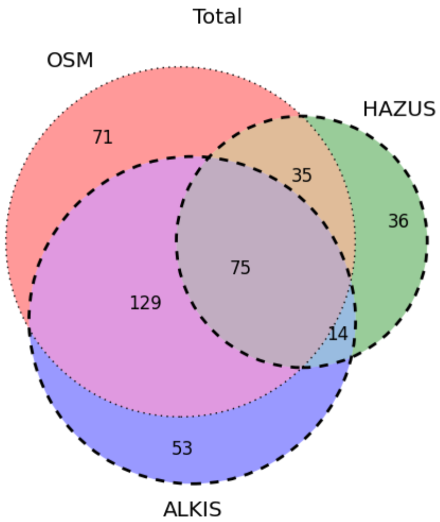

Figure 3. Venn diagram of asset types for different reference catalogues

Figure 3 shows that the HAZUS and ALKIS reference catalogues contain different asset types to a great extent. The HAZUS reference shares about one third of all asset types listed in ALKIS, while vice versa the ALKIS reference shares about 55% of all asset types listed in HAZUS. On the other hand OSM covers and about 69% of all asset types from HAZUS and about 75% of all asset types from ALKIS.

Further analysis shows that OSM bears great potential to add asset types to the reviewed reference catalogues. A combined list of asset types from OSM and ALKIS will cover about 78% of HAZUS. A combined list of asset types from OSM and HAZUS will cover even 82% of ALKIS. Beyond that, there are 71 asset types that can neither be found in HAZUS nor in ALKIS, but in OSM (e.g. cable distribution cabinet, road under construction, automated external defibrillator, fire extinguisher).

FINAL REMARKS AND FUTURE WORK

This paper presents first results towards a method for assessing the quality of OpenStreetMap for the specific purpose of identifying assets of critical infrastructure in support of emergency planning. First results show that critical asset types of all selected sectors and branches are well represented in OSM.

OSM provides good results for most of the critical sectors compared to other reference catalogues from HAZUS and ALKIS. The high number of asset types in OSM shows that the OSM data structure is suitable to represent critical infrastructures in a sophisticated way. In total, OSM reaches a level of detail that is comparable to other reference catalogues.

However, regarding different critical sectors the results also show that other reference catalogues perform better than OSM in specific areas. Especially in the sector “Health”, HAZUS contains much more asset types. In this manner, it should be noticed that using OSM as an exclusive source will exclude a significant number of asset types of critical infrastructures.

However, our results show that OSM bears great potential to add a high number of asset types to both HAZUS and ALKIS reference catalogues.

Further research should build upon these first results to assess the fitness for use of the OSM data for identifying assets of critical infrastructure. For doing this, the data that is actually produced by volunteers should be analyzed for different case studies, in order to address the local heterogeneity in the quality and completeness of the OSM data. This is crucial to validate our hypothesis out of the scope of our analysis (Germany and USA) and for filtering incorrect data.

Furthermore, if the quality of the data is confirmed, we envisage the design of a decision-support system that would be able to retrieve data automatically from OSM and present it to the user for assisting emergency planning, based on the catalogue presented in this paper.

ACKNOWLEGDEMENTS

The authors are thankful to Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW) for providing funding for this research within the "Klimopass" program. The authors would like to thank Svend-Jonas Schelhorn for his support regarding the analysis of the reference catalogues. J. P. Albuquerque is grateful for CAPES (grant no. 12065-13-7) and Heidelberg University (Excellence Initiative II / Action 7) for supporting his contribution to this research.

REFERENCES

1. Arsanjani, J., Barron, C., Bakillah, M., and Helbich, M. (2013) Assessing the Quality of OpenStreetMap Contributors together with their Contributions. Proceedings of the 16th AGILE conference, Springer, Leuven.
2. Australian Government. 2010. Critical Infrastructure Resilienz Strategy. Online: www.tisn.gov.au/Documents/Australian+Government+s+
3. Critical+Infrastructure+Resilience+Strategy.pdf Acess: January 30, 2015
4. Barron, C., Neis, P., and Zipf, A. (2013) A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. Transactions in GIS, 8, 6, 877–895.
5. Bender, S. (1991) Primer on natural hazard management in integrated regional development planning. Organization of American States, Department of Regional Development and Environment. Executive Secretariat for Economic and Social Affairs, Washington, DC.
6. Bouchon, S. (2006). The Vulnerability of interdependent Critical Infrastructures Systems: Epistemological and Conceptual State of the Art. Institute for the Protection and Security of the Citizen, Joint Research Centre, European Commission, 99.
7. Bundesministerium des Innern. Nationale Strategie zum Schutz kritischer Infrastrukturen. Online: http://www.bmi.bund.de/cae/servlet/contentblob/544770/publicationFile/27031/kritis.pdf Acess: January 30, 2015
8. Bush, George W. 2003. The national strategy for the physical protection of critical infrastructures and key assets. Online: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA413033. Acess: January 30, 2015
9. European Union. 2009. The identification and designation of European critical infrastructure and the assessment of the need to improve their protection. Online: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:345:0075:0082:EN:PDF Acess: January 30, 2015
10. Haklay, M. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and planning. B, Planning & design, 37, 4, 682–703.
11. Hecht, R., Kunze, C., and Hahmann, S. (2013) Measuring completeness of building footprints in OpenStreetMap over space and time. ISPRS International Journal of Geoinformation, 2, 4, 1066–1091.
12. Moteff, J., & Parfomak, P. (2004). Critical infrastructure and key assets: definition and identification. Library of Congress, Washington DC Research Service.
13. Mooney, P., & Corcoran, P. (2012). The Annotation Process in OpenStreetMap. Transactions in GIS, 16(4), 561–579. doi:10.1111/j.1467-9671.2012.01306.x
14. Neis, P., Singler, P., and Zipf, A. (2010) Collaborative mapping and Emergency Routing for Disaster Logistics–Case studies from the Haiti earthquake and the UN portal for Afrika, Geoinformatics Forum Salzburg,

239–248.

15. Organization of American States (OAS) Department of Regional Development and Environment (DRDE). 1991. Primer on Natural Hazard Management in Integrated Regional Development Planning, 440. Department of Regional Development and Environment, Washington D.C.
16. Parliament, E. 2007. Directive 2007/60/ec of the European Parliament and of the council of 23 October 2007 on the assessment and management of flood risks. Official Journal of the European Union (2455) 27–34.
17. Renda, A., & Hammerli, B. (2010). Protecting Critical Infrastructure in the EU. CEPS Task Force Report.
18. Schelhorn, S. J., Herfort, B., Leiner, R., Zipf, A., & Albuquerque, J. P. de. (2014). Identifying Elements at Risk from OpenStreetMap: The Case of Flooding. In S. R. Hiltz, M. S. Pfaff, L. Plotnick, & P. C. Shih (Eds.), Proceedings of the 11th International ISCRAM Conference (pp. 508–512). University Park, Pennsylvania, USA,: ISCRAM.
19. Soden, R., Budhathoki, N., and Palen, L. (2014) Resilience-Building and the Crisis Informatics Agenda: Lessons Learned from Open Cities Kathmandu. Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management (ISCRAM) 2014, State College, Pennsylvania.
20. Turoff, M., Hiltz, S. R., Bañuls, V. A., and Van Den Eede, G. (2013) Multiple perspectives on planning for emergencies: An introduction to the special issue on planning and foresight for emergency preparedness and management, Technological Forecasting and Social Change, 80, 9, 1647–1656.