The Inevitability of Calibration in VGI Quality Assessment

Franz-Benjamin Mocnik
Heidelberg University
GIScience Group
Im Neuenheimer Feld 348
69120 Heidelberg, Germany
mocnik@uni-heidelberg.de

Alexander Zipf
Heidelberg University
GIScience Group
Im Neuenheimer Feld 368
69120 Heidelberg, Germany
zipf@uni-heidelberg.de

Hongchao Fan
Heidelberg University
GIScience Group
Im Neuenheimer Feld 348
69120 Heidelberg, Germany
hongchao.fan@uni-heidelberg.de

Abstract

Data and the context in which data is interpreted are subject to imperfection, and the interpretation of the data accordingly depends on the choice of the context. Data quality and fitness for purpose can thus not be assessed without any choice of a context, a fact that can be regarded as an inevitable calibration of the data quality assessment. The paper examines this effect, why Volunteered Geographic Information is particularly prone to such influences of the choice of a context, and how this influence on the data quality assessment can possibly be reduced. Different types of contexts are discussed at the example of OpenStreetMap data.

Keywords: Data quality; fitness for purpose; calibration; OpenStreetMap (OSM); Volunteered Geographic Information (VGI).

1 Introduction

Volunteered Geographic Information (VGI) is created in a very heterogeneous process. Volunteers are representing the environment based on different perception methods, having local or remote knowledge, and with differing previous knowledge and cultural background. Quality assessment is accordingly very important in case of VGI, and many quality indicators have been discussed. Existing indicators use different contextual information to assess the quality of the data, that is, other information, to which the original data can be compared. Amongst these contexts are spatial and temporal information as wells as principles. A comparison of data to such a context favours certain interpretations, and these interpretations may reveal information about the quality of the data. If the data are examined in different contexts, it is yet not clear how to merge possibly conflicting findings about data quality.

The paper intends to inquire the foundations of context and groundings in data quality assessment, with the aim to argue the need and inevitability to calibrate in quality assessment. It does not intend to discuss practical examples. The role of contexts and groundings in the creation process of VGI data is examined in Section 2. The process of using the data can, to some degree, be regarded as being opposed to the process of the data creation, as is discussed in Section 3. Data use accordingly depends on contexts and groundings as well. Section 4 discusses the differences between context in VGI and non-VGI. Section 5 provides examples of contexts that can be used for the assessment of the quality of OpenStreetMap (OSM) data. Finally, it is discussed in Section 6 that data quality assessment depends on contexts and groundings as well. This dependency can be regarded as a calibration and can be expected to be reduced by the combination of different contexts.

2 Creation of VGI data

The creation and use of VGI is a complex process that is shaped by a community of volunteers, each of them creating and using the data in different contexts. The creation of data in general, and VGI data in particular, starts with the perception of the environment, for example, by the visual perception of streets and houses, by the auditory perception of cars on the streets, or by examining aerial images, which have been produced by measurement instruments mounted on satellites or airplanes. Even the interaction with the environment can be seen as a perception because interactions are able to reveal properties of the environment.

The perception of the environment results in a mental representation, which provides information about the properties and the interaction possibilities of the environment. Such a mental representation can, however, not be accessed by other people, and we thus produce formal data, which can be shared and be interpreted by others. As an example, the OSM database contains a formal representation of the environment in means of nodes, ways, and relations. These elements have been created by different users, having different mental models of the environment in mind while referring to the, for the most part, commonly accepted folksonomy of OSM.

There exist many aspects of context besides the context given by the perception of the environment, amongst them the context of the perception method (visual, auditive, using technical instruments like mobile phones or aerial images, etc.); previously gained knowledge; the context rendered by the community, in case of VGI; and the cultural context. These contextual aspects are of particular interest, because they render the understanding of how the data can be interpreted, and how the formal symbols (for example, characters, words, or numbers) of a dataset could refer to the environment. If data get a meaning by being aware of this context and by thus knowing how they refer to the environment, the data are *grounded* in the environment. This process of grounding has been identified as a solution of the *problem of reference* by relating symbols to perceptual operations (Scheider, 2012). Information and its quality are accordingly rendered by the context and the grounding.

Data creation and data use have similarities because both relate the environment to formal data. These processes point, however, in opposing directions: data creation starts with the environment, while data use ends in actions in the environment. We have discussed the dependency of the data creation process on context and grounding in this section. The next section examines the same dependency for the process of data use.

3 Fitness for purpose depends on context

Data can be used for many different purposes. OSM data are, for example, used for routing and navigation tasks, as well as for geocoders and for producing maps for different kinds of usage scenarios. All these usages suggest possible interpretations of the data, and the interpretations even depend on many additional factors, among them the cultural background of the user. Different interpretations of a dataset can not only result in differing information but also in contradicting information. The interpretation of a map as a cyclist or as a driver of a motor vehicle may, for example, lead to different information about distances. This demonstrates that the context in which the data are grounded is also essential for the use of the data. Similar considerations have been made in multi-agent simulations of map creation and use (Frank, 2000).

Data can only successfully be used if the user is able to ground the data by finding a context that affords a suitable interpretation of the data. This interpretation needs to reflect the context of the creation of the data, as well as the context in which the data should be used. A map produced by OSM data can, for example, only be used for routing tasks, if the reader of the map is able to interpret the map – this interpretation depends on the creation of the map – and if the reader is able to understand how routing tasks relate to his interpretation of the map – this understanding depends on the context of the navigation task. This possibility of the data to be interpreted in a suitable context can be understood as a feature of the data. Such a feature of an object – in this case the data – to be used in a certain way in a suitable context is called an affordance (Gibson, 1977; Sanders, 1997; Turvey, 1992).

Fitness for purpose describes how well data can be used for a certain purpose. This in turn depends on whether a suitable interpretation of the data exists, that is, whether a suitable context and a suitable grounding exist. If no suitable context can be found, or possible interpretations of the data are not known, the data cannot be used. It is not the data themselves that the fitness for purpose refers to, nor is it the data in a given or all contexts. Instead the fitness for purpose refers to whether a suitable context exists, that is, it refers to an affordance of the data. Consider, for example, a map of NYC. The map does not afford to perform routing tasks in the area of Chicago, even though the map is fit for the purpose of routing tasks because it renders routing tasks in the area of NYC possible.

4 Differences between the different contexts (and groundings) in VGI and non-VGI

Contexts and groundings are, in principle, very similar in VGI and non-VGI data in the sense that both can be defined in the same way. There are, however, many practical differences. As VGI is, in contrast to non-VGI, created by many people, many contexts and groundings exist simultaneously. This is in contrast to non-VGI data, which is created by a small group of people or even only one person, often using only one perception method and having a fixed taxonomy or ontology in mind. When data is created while having only one context in mind, this context can be considered as a natural choice for the interpretation of the data. In case of VGI, there does not exist such a natural choice. Instead, it would suggest itself to consider several contexts at once for the interpretation, namely the contexts that have been used during the creation of the data.

The interpretation of OSM data mostly happens in some given context. Consider, for example, a map of a town. It may serve for many purposes, possibly in several contexts defined by these purposes, but when a purpose has been chosen, we usually do not change this context when considering different features of the map. It is yet unclear, why this strategy of assuming only one context works for OSM and VGI in general. Possible explanations may be that the contexts are much more similar than one might expect, and that several contexts can be merged in a meaningful way, resulting in more general concepts. Also, the assessment of the data quality and the fitness for purpose have to take care of several contexts in case of VGI. At the same time, the assessment has to take into account that the data will, at the end, very likely be interpreted in only one or very few contexts, and that the context will most likely not vary for every feature of the map.

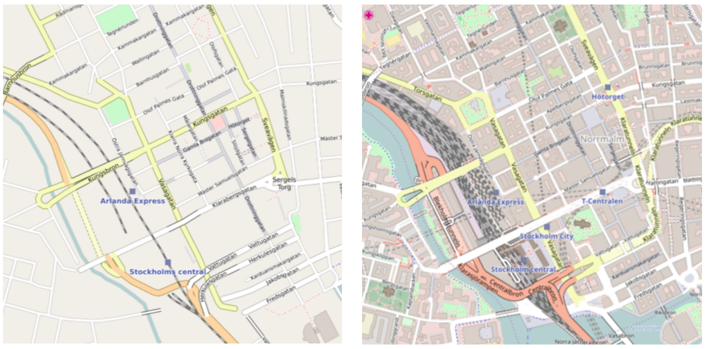
5 Examples of contexts in quality assessment

Quality assessment and the fitness for purpose can be assessed in different contexts. In this section, we exemplify the use of different contexts for the quality assessment of OSM data: the context of principles, a historic context, and a spatial context. The discussed examples are only presented very shortly and are not discussed in detail, as they are only meant to illustrate the impact of different contexts in quality assessment. Real-world scenarios require more sophisticated considerations.

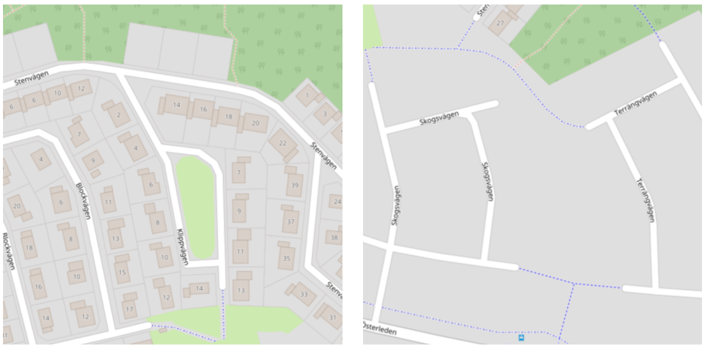
Context of principles. Our perception of the environment is often memorized, resulting in knowledge about the environment. Such knowledge reveals certain patterns, structures, and principles, because the environment is, amongst others, shaped by the laws of nature and by human activity patterns. The concept of streets and crossings is, for example, widely accepted and used – even if the concepts strongly depend on context and cultural background – and the perception of streets and crossings do thus not come unexpectedly. Buildings are often located near a street; streets shape connected networks; and the street network often consists of layers, each of these layers containing streets of a certain type; etc. These principles are, unlike natural laws, often of heuristic nature. They are not universal but rather apply often while clear counter examples exist. There are buildings far away from any street, and there are streets that

Figure 1: Examples of contexts. (a) Context of principles: we would expect the street network to be connected, but the street aligning with buildings no 5 and 6 is unconnected to the other streets; this may indicate a street to be missing in the map, (b) historic context: the older map (left) seems to contain much less detail than the newer one (right); this indicates that the newer map is more accurate, and (c) spatial context: the number of houses represented in the data is very different despite of the very similar street layout, even in the same town; this may indicate that existing houses are not represented in the right map. Maps © OpenStreetMap contributors (cf. http://openstreetmap.org)

(a) Street network in Kiruna, Sweden



(b) City centre of Stockhom, mapped in 2007 (left) and 2017 (right)



(c) Two areas in Kiruna

are not connected to any street network. When assuming that a principle holds in general, one can identify unexpected situations described in the dataset. In the context of the principle of 'connected street networks', one may identify a street not connected to the street network as being a situation that does, in all likelihood, not exist in reality. In case that many of such unconnected streets exist, one may, for example, conclude, that it is very probable that the street network represented in the OSM data is being far from complete.

Historic context. The history of OSM data reveals information about its creation process and the mapping behaviour. The completeness of OSM data can accordingly be assumed to be reflected by the history of the data: unless all streets in a certain area have been mapped, streets not yet represented in the OSM data will continuously be mapped, and the number of streets represented in the data increases. If this process does not stop before all streets are represented in the data, an assumption which is true in most cases, the number of streets represented in OSM data converges to the number of streets existing in the environment. This *saturation* effect has been discussed in literature (Neis, et al., 2012; Gröchenig, et al., 2014; Barron, et al., 2014), in particular for streets – objects tagged as *highways* in the OSM database.

Spatial context. The arrangement and number of buildings and of streets is, for example, not completely random. A city with 100,000 houses is likely to have a higher number and a higher density of streets than a village with only 100 houses, and the arrangement of streets in a large city exposes characteristic differences compared to a small village (Hagenauer & Helbich, 2012; Fan, et al., 2014). If a street network similar to the one in Oxford is contained in the OSM data, one may assume that the density of houses in the considered area is about as high as in Oxford. A dense street network without any house can, for example, be interpreted as being an indicator for an incomplete mapping of houses.

None of the contexts and corresponding examples discussed in this section is able to *prove* statements about data quality and fitness for purpose, because the considerations are of heuristic nature. They rather are able to suggest certain interpretations of the data, and a comparison of the data in different contexts might even lead to stronger interpretations. In the next section, we will discuss this effect of different interpretations due to the heuristic nature in more detail.

6 Calibration of data quality indicators

Context and grounding are imperfect. Otherwise, there would be no need to assess data quality, and imperfection is, in fact, an inherent property of complex spatial data (Couclelis, 2003). The interpretation of data relies on the context and the grounding of the data in the environment, as has been discussed in Section 2 and 3. Different interpretations can, accordingly, result in differing, sometimes even contradicting conclusions. In the context of a motorist, for example, certain features, such as the cycling distance between two locations, remain hidden and can hardly be understood. Compared to other data, VGI is especially prone to such imperfection due to the heterogeneity of contexts, as has been discussed in Section 4. This heterogeneity implies that the data cannot be interpreted in *one* 'ideal' context. Instead, a deeper comprehension is needed of how the resulting VGI data reflect this heterogeneity, and how quality assessment can contribute by filtering and modifying the data such that they can be interpreted in at least *one* meaningful context.

The imperfection of the context and the grounding has consequences for the quality assessment. If contexts and groundings would not be subject to imperfection, the examples of contexts in Section 5 would all lead to the same judgement of the data quality, and data quality would be independent of such a choice. The choice of a grounding becomes, however, relevant for real-world scenarios, because the imperfection can render differing results. Assume, for example, that the completeness of buildings in a town is to be assessed by the spatial context of another town as a reference. If the buildings in this reference town are themselves far from complete, the incompleteness of the buildings in the town to assess cannot be uncovered. This incompleteness may, however, be obvious when assessing against the context of principles: there may in both cities be no buildings despite of a very dense street network. This demonstrates that the imperfection implies the need to choose a context and a grounding to assess data quality, a procedure which can be seen as a calibration: once a context chosen, the quality of different parts of the data can be assessed against the chosen context.

The calibration of the process of data quality assessment is not only needed, in case of VGI, but the data quality assessment is also inevitable subject to calibration. There is no 'natural' context, defined by the creation of the data, which suggests itself. Instead, one or more contexts have to be chosen – data quality cannot be assessed without.

The assessment of data quality and fitness for purpose should ideally be independent of any choice of a context, because also the data quality and the fitness for purpose themselves are. It is hard to conclude any 'objective' finding, if different choices of a context or a grounding result in different findings about data quality. Only the context and the grounding can be chosen, in case some data should be assessed; the data to assess is however fixed. Such a choice of the context is thus an asymmetric situation. While the calibration and asymmetric situations in general cannot be avoided, several asymmetric situations can be assessed and the resulting findings can be integrated. In the best case, an asymmetric situation can be transformed into a symmetric one by considering all combinations. When a city is, for example, to be assessed against the context of another city, cities can mutually be compared, which results in a symmetric situation. It can be hoped for that the impacts of the choices are compensating in the best case, and that the resulting findings about data quality depend less on the choices and expose less bias.

7 Conclusion

We have discussed how the context and the grounding impact the assessment of data quality and fitness for purpose. A major reason for this impact is the effect of the mental model, amongst others influenced by the cultural background of the person who creates the data. Future research may consider the influence of cultural background in greater detail.

The imperfection has been argued to render the need and the inevitability of calibration in the assessment of data quality. This inevitability can possibly be countered by combining different quality indicators. Future research may understand how such strategies can improve the quality assessment, and practical examples have to be examined. Besides the need to calibrate, the imperfection also affects the consistency, as well as the credibility of the data quality assessment. Both may be examined in greater detail.

Acknowledgement

This work has been partially supported by the DFG project *A framework for measuring the fitness for purpose of OpenStreetMap data based on intrinsic quality indicators* (FA 1189/3-1).

References

Christopher Barron, Pascal Neis, and Alexander Zipf. A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Transactions in GIS*, 18(6):877–895, 2014.

Helen Couclelis. The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS*, 7(2):165–175, 2003.

Hongchao Fan, Alexander Zipf, and Qing Fu. Estimation of building types on OpenStreetMap based on urban morphology analysis. *Proceedings of the 17th AGILE Conference on Geographic Information Science*, page 19–35, 2014.

Andrew U. Frank. Spatial communication with maps: defining the correctness of maps using a multi-agent simulation. *Proceedings of the International Workshop on Maps and Diagrammatical Representations of the Environment*, page 80–99, 2000.

James J. Gibson. The theory of affordances. In Robert Shaw and John Bransford, editors, *Perceiving, acting, and knowing*, page 127–143. Lawrence Erlbaum Associates, Hillsdale, NJ, 1977.

Simon Gröchenig, Richard Brunauer, and Karl Rehrl. Estimating completeness of VGI datasets by analyzing community activity over time periods. *Proceedings of the 17th AGILE Conference on Geographic Information Science*, page 3–18, 2014.

Julian Hagenauer and Marco Helbich. Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. *International Journal of Geographical Information Science*, 26(6):963–982, 2012.

Pascal Neis, Dennis Zielstra, and Alexander Zipf. The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet*, 4:1–21, 2012.

John T. Sanders. An ontology of affordances. *Ecological Psychology*, 9(1):97–112, 1997.

Simon Scheider. *Grounding geographic information in perceptual operations*. IOS Press, Amsterdam, 2012.

Michael T. Turvey. Affordance and prospective control: An outline of the ontology. *Ecological Psychology*, 4(3):173–187, 1992.