RESEARCH ARTICLE

A grounding-based ontology of data quality measures

Franz-Benjamin Mocnik, Amin Mobasheri, Luisa Griesbaum, Melanie Eckle, Clemens Jacobs, and Carolin Klonner

GIScience Research Group, Heidelberg University, Germany

Abstract: Data quality and fitness for purpose can be assessed by data quality measures. Existing ontologies of data quality dimensions reflect, among others, which aspects of data quality are assessed and the mechanisms that lead to poor data quality. An understanding of which source of information is used to judge about data quality and fitness for purpose is, however, lacking. This article introduces an ontology of data quality measures by their grounding, that is, the source of information to which the data is compared to in order to assess their quality. The ontology is exemplified with several examples of volunteered geographic information (VGI), while also applying to other geographical data and data in general. An evaluation of the ontology in the context of data quality measures for OpenStreetMap (OSM) data, a well-known example of VGI, provides insights about which types of quality measures for OSM data have and which have not yet been considered in literature.

Keywords: data quality, fitness for purpose, data quality measure, grounding, ontology, geographical data, OpenStreetMap, osm, volunteered geographic information, vgi

1 Introduction

The complexity of our environment cannot be completely represented in a dataset, and representations are often limited to certain aspects. Accordingly, data is often created with having a purpose in mind: What is the desired scope of the data? Which aspects are of importance? How fine-grained does the representation need to be? Which data retrieval methods are available? and so forth. Not only data quality but also the fitness for a purpose is thus examined: Are we able to successfully use data for a certain purpose? The issues of data quality and fitness for purpose apply not only to datasets but also to information

that is shared visually or orally. Information being of poor quality or information not being fit for a certain purpose may have far-reaching consequences, as in the case of flight *CRX 3597*, which crashed on a hill near Zürich airport on the 24th of November 2001, due to incomplete information about the weather and about the hill among other things [60]. This demonstrates the relevance and the importance of data quality and fitness for purpose.

The issue of data quality receives increasing attention, both in science and in industry. One reason might be the availability of large and increasingly heterogeneous datasets. Among the examples of such heterogenous data is volunteered geographic information (VGI) that is created by the voluntary effort of groups of people who collaborate, for example, in OpenStreetMap (OSM), and ambient geographic information (AGI) resulting from people sharing information without the intention to contribute geographical information to a larger dataset or data collection, for example, Twitter and Flickr. People contributing to VGI and AGI usually have very different aims and different abilities, which is in contrast to conventional datasets that are often created by a few experts who follow strict rules and guidelines. The heterogeneity of the contributors is reflected in the heterogeneity of the data. Rather, the quality of such heterogeneous datasets can accordingly not be assessed in its entirety by the interpretation of the data with a fixed ontology and a subsequent comparison to reference data. A better understanding of the processes that generate the data is required to assess data quality, which provides insights into how the data can be interpreted—there may not even exist *one* fixed ontology for the entire dataset but the interpretation may happen on an ad hoc basis. The *internal* characteristics of a heterogeneous dataset seem thus to be important with respect to data quality. Among these characteristics are the process that generated the data, the history of the data, the methods to retrieve the data, the sources from which the data is aggregated, and the different purposes for which the data has been collected.

Numerous measures of data quality have been introduced and discussed in the literature. Such measures are methods to determine, for a given dataset, the quality of the dataset. One obvious choice is to compare, for example, a map to the actual environment or to a reference dataset that is assumed to be of high quality. On the other side, also the internal characteristics of a dataset can be used to assess its quality, for example, the growth of a street network over time [5]. As a distinction between "approaches for intrinsic assessment [. . .] that do not use external reference data" [42] and intrinsic measures that can be evaluated "without the usage of any reference data" [5] the concepts of *intrinsic* and *extrinsic* measures have been introduced. While the authors are not aware of any formal definition of intrinsic and extrinsic measures in literature, the distinction between both seems to be manifest: if a dataset shall be assessed for its quality by a certain measure, this measure is called intrinsic if it does not require any additional data besides the dataset, otherwise it is called extrinsic. This definition only makes sense if the measure is to be characterized in combination with the dataset, because the same measure may be able to operate on different datasets as well as on one dataset only. The completeness of buildings in a city in OSM data might, for example, be assessed by comparing the specific city in OSM data to other cities in cadastres, or by comparing it to other cities within OSM data in order to gain an understanding of the structure of cities and thus of the number and distribution of buildings to be expected. Exactly the same methodology and the same measure might be used in both cases, because the structure of cities in general can be studied from cadastres as well as from OSM data. In the case of other cities within OSM data, the measure would be intrinsic while being extrinsic in the second case. This demonstrates that there is a need to characterize data

quality measures themselves in more detail, despite the usefulness of the distinction between intrinsic and extrinsic measures.

One widely used classification scheme is provided by the International Organization for Standardization [39]. Measures are classified by this scheme according to several aspects of data quality such as positional accuracy or completeness. This classification refers to *what* is measured. In contrast, the classification into extrinsic and intrinsic measures refers to *which* data is used by the measure—data from the same dataset or other data. In this article, we extend this question to *where* the information comes from that allows us to draw conclusions about data quality. Imagine the following thought experiment: A dataset claims that a tree T is located inside a building B. It cannot be concluded by the statement "T is inside B" itself whether this very statement is correct and agrees with reality. In fact, trees are usually not located in buildings, and buildings do not usually accommodate trees. Counterexamples exist though, for instance, trees in entrance halls of major insurance companies. Additional information about how far this statement correctly describes the real environment (i.e., about the external quality of the above statement) can only be concluded on the basis of our perception of the environment or on the basis of further information that itself refers to the environment. Thus, the *where* question from above refers to the source from which additional information is gained in order to be used by the measure. Sources for such additional information can be our perception of the environment, heuristic principles, other data from the same dataset, etc.

This article discusses to which sources of information data can be compared when assessing the quality of the data. By classifying quality measures according to these sources, the article provides a systematic view on the assessment of data quality and fitness for purpose. After a short discussion of existing ontologies and taxonomies related to quality measures and data quality, as well as of other related work (Section 2), the notion of grounding is introduced in the context of data quality (Section 3). It is argued that data quality can be measured by comparing different groundings. We accordingly introduce a grounding-based ontology of quality measures (Section 4). While the ontology applies to all kinds of data, we investigate the ontology for geographical information. Two examples of VGI are provided (Section 5). Finally, the ontology is evaluated in the context of quality measures for OSM data (Section 6).

2   Related work

The uncertainty of data is among the topics that frequently reoccur in many contexts of GIScience [7, 29, 30]. A major part of research about the uncertainty of geospatial data concerns the understanding of data quality and fitness for purpose [62]. These concepts have been addressed by comparison to ground truth, which is, among others, discussed in ISO 19157 [39], as well as by understanding the use of data [12, 24, 26]—two perspectives of the same problem [17]. It is of high importance to ensure the quality and fitness for purpose of geographical information in order to minimize errors and strengthen the reliability of data usage, regardless of their nature and of how they are collected [59]. Limiting factors of uncertainty are, however, an inherent property of geographical information and knowledge, and these limitations cannot be evaded [14].

Several ontologies and taxonomies have been introduced in order to formally represent aspects of data quality and fitness purpose. We discuss the most important ones in the

remainder of this section, with the aim to relate them to the grounding-based ontology proposed in this article. Frank [25] has proposed an ontology about which commitments, made during the collection and aggregation of information, cause imperfect knowledge. These commitments refer, for example, to the limitations caused by partial knowledge, to observation errors, to simplification and classification, as well as to the context. While the ontology describes the process that leads to imperfect knowledge, and hence also the reason for why knowledge is imperfect, it does not capture how data quality can be assessed.

In contrast to Frank's ontology, several data quality dimensions have been discussed in literature—aspects of data that indicate good or poor quality. Possibly the first thorough discussion of data quality dimensions that separates information from the real world it represents was published by Wand and Wang [64]. They discuss completeness, unambiguity, meaningfulness and correctness as intrinsic data quality dimensions. In addition, they summarize data quality dimensions that have been discussed in literature, including accuracy, reliability, timeliness, and many others. The International Organization for Standardization has listed "data quality elements" for geographical information, which correspond to previously published data quality dimensions [39]. Most of these data quality elements refer to a comparison to ground truth, such as positional and thematic accuracy, but also other dimensions are listed, for example, the usability. Data quality dimensions often serve as categories that guide the discussion of data quality. They can be used to characterize the quality of data, and they can be used to characterize data quality measures by the dimensions they measure. Quality dimensions cannot, however, be used to characterize quality measures by *how* they measure and *which* information they rely on. This is what this article aims at.

Data quality can be improved or ensured even without being explicitly measured. In the case of VGI datasets, three categories of mechanisms to ensure data quality have been proposed [31]. The category of *crowd-sourcing approaches* consists of approaches to let individuals of a community validate data and correct errors. Observations of animal or plant species can, for example, be strengthened when other volunteers make similar observations in a nearby area [41]. Such crowd-sourced approaches can be distinguished from *social approaches*. As the result of a social process, trusted individuals, who have made themselves a good reputation with their contributions to VGI or have prior expert knowledge, maintain and control the quality of other volunteers' contributions [43]. The comparison of VGI data, which has been produced by volunteers, to data that has been created by selected experts with local knowledge about the mapped area can be seen as a typical example [18]. Finally, as a *geographic approach*, VGI can be compared to existing geographical knowledge, among them scientific laws and existing datasets. The total length of road features compared to the road length in reference data has, for example, been used for measuring the completeness of OpenStreetMap (OSM) [33], and a statistical measure for identifying spatial patterns in the occurrence of flood-related tweets associated with proximity to and severity of flood events has been introduced [16]. *Data mining approaches* have been independently proposed as another category [58]. The use of machine learning has, for example, been demonstrated to detect and improve inappropriate classification of OSM features [1]. Many more examples for each of these categories exist. While this classification is able to categorize mechanisms for ensuring data quality, it does not provide answers to which information data quality measures can rely on.

Bordogna et al. [8] have categorized methods to ensure data quality into *ex ante* and *ex post* approaches. Methods categorized as ex ante ensure data quality before the data is

created or even collected. Typical approaches incorporate the education of contributors, for example, to measure positions with better accuracy or to recognize certain types of objects on aerial imageries [44], and the stimulation of discussions in the community that lead to a more consistent conceptualization. These methods differ from ex post approaches, which improve data quality after the data has been collected. As an example, VGI data is often maintained by experienced users that have a good reputation [43]. The categorization into ex ante and ex post approaches complements Goodchild's categorization of quality ensuring methods. The distinction explains when the approach is used, and accordingly, which information is already available. However, it does not explain to which source of information the approach refers, and how the approach is grounded.

An overview over different ontologies about data quality in the case of VGI has been provided by Fonte et al. [22]. As we have discussed, existing ontologies point out which aspects of data quality exist, how data quality measures can be categorized by these aspects, which commitments lead to imperfection of knowledge, and how data quality can be ensured. To our knowledge however, there exists no literature providing a structural view on which information is used by quality measures in order to assess the quality. This is despite the fact that a measure can be characterized by the used information and the way it relates to the physical environment. A measure that explicitly builds on a comparison to our perception of the environment is relatively independent of a purpose. In contrast, a measure referring to the data that is to be assessed, or even to the data after being processed, depends, at least in parts, on the purpose for which the data has been created. In addition to this dimension of incorporating the purpose, measures can be characterized by whether they use data or whether they use rules derived by the data. In the next section, we discuss more thoroughly why the question of which information is used by a measure is of high importance for the understanding of how quality measures work. This understanding will finally render possible the introduction of a meaningful and original ontology of data quality measures by which information is used.

3   Grounding quality measures

Data quality measures describe how well we are able to relate the data to the environment. Is the data logically consistent and usable? That is, are we able to interpret the data? Is the data complete, the contained locations and the thematic information of high accuracy, and the temporal information of high quality, etc.? Quality measures provide answers to these questions by either relating the internal properties of the data to other data or to the physical environment (e.g., in the case of spatial and thematic accuracy) or by relating to the data itself (as in the case of consistency.) In the first case, the data needs to be related in some way to the environment by understanding how the data refers to physical objects, their features, and processes of the environment. Logical consistency can in the second case be assessed without any need to relate the data to the environment, but it refers to how far we are able to make sense of the data by relating it to the environment—data that is logically inconsistent cannot easily be related to the environment in a meaningful way. In particular in the first case, quality measures relate the data to the environment, which raises the question of *how* measures achieve such relations. This section discusses the process of establishing such a relation in more detail.

Physical objects and processes incorporating physical objects can be "experienced" by the interaction with them and by the perception of the results of such interactions by our human senses. We refer, in the scope of the article, to these objects and processes as our *environment*, because we are physically embedded in the environment and are able to closely interact. This definition of an environment shall not imply any ontological commitment: the environment exists of everything that is not the direct product of our cognition, which is in contrast to information because it is created by our cognition. Two types of interactions with the environment can be distinguished: *actions* in the environment, which affect the environment by modifying it, and the *perception* of the environment, by observing objects and processes while gaining an understanding of potential activities within the environment [47, 53, 54]. Besides human perception by the use of our senses, perception can also incorporate technical sensors and machines—for example, sensors and machines transform the environment temperature or the location of the sensor into electrical current, which then can be observed and interpreted by humans. The term perception thus refers, at least in the scope of this article, to various kinds of processes that create data out of the environment. Although there exists no clear distinction between actions and perceptions in general, both can be prototypically illustrated by the following example. When we are in a town we see roads, street signs, etc. We perceive the environment and hence an understanding of the street network emerges. The gained understanding can at least in parts be formally represented as a graph, as is done in OSM. This graph can be used to perform actions in the environment, for example, to navigate from location A to location B when reading a map or using a routing service. This navigation takes place in the physical environment, and we find ourselves being relocated to location B. This is a change of the environment that we are, as physical objects, a part of. Our understanding of us being able to perform a certain action in the environment is often referred to as an *affordance* of the environment [27, 56], that is, a potential activity, because it is the environment which affords the action.

Data refers to the environment when being interpreted. Formal data, like meteorological data stored in a spreadsheet, datasets used to produce maps, or spatial information written down by hand, consist of formal symbols, which are arranged in different ways. These symbols have no meaning unless they get interpreted: we give a meaning to these symbols by explaining how they refer to entities and features of the environment. An element in OSM data with the tag `"highway"="residential"` is, for example, commonly interpreted as representing a residential street, which is a physical object of the environment. This interpretation refers to our perception of the environment and its affordances: we know what a street looks like according to our concept of a street, and that we are able to walk or drive on the street. This process of relating data to the environment by perception and affordances [27] has been termed *grounding* [35]. The process of grounding data is not a formal process and hence cannot be captured by our ratio in its entirety. Instead, a grounding always incorporates perception of the environment and its affordances.

There usually exists more than one possibility to interpret symbols, which is why misunderstandings arise: Which measurement uncertainty does the value "286 K" expose? What is the exact definition of the category of a street? Should a certain direction be referred to as "left" or "left ahead"? On the other side, interpretations are not random, and some interpretations are to be favored. There is no interpretation of meteorological data as being a representation of a street network that would suggest itself, even if both can be represented as tables or text. Rather, we are able to identify meteorological data as incorporating temperature or humidity just by the way symbols internally relate inside the dataset; we are able to

identify numbers, judge their meaning from their range; and we may be able to guess that twelve values being separated by commas is a list with one value per month. The problem of finding such favored interpretations by the internal characteristics of the data has been termed *symbol grounding problem* [35]. It has been discussed for geographical information by Scheider [56].

Data quality is rendered by the imperfection of groundings. After the creation of data in a process of perceiving the environment, incorporating previous knowledge, and formally representing the information, the resulting data is grounded by this very process of data creation: we know how the data refers to the environment because of our knowledge about their creation process. The imperfection of this process, including the symbol grounding problem, results in the data being incomplete, being subject to uncertainty, and suffering from other issues related to data quality. This lack of data quality cannot, however, be uncovered by grounding the data anew using the same or a very similar perception or previous knowledge, because this would not add any new facts to the assessment—the new grounding would suffer from the same imperfection. Reading the same thermometer several times in exactly the same way will, for example, only result in a duplicate value, but another thermometer would be needed to understand how well the resulting temperature value describes the temperature of the environment. The value "286 K" might be grounded by our knowledge that a certain thermometer has been used, but it might stay unclear under which conditions the measurement was performed, for example, with or without any windshield. Accordingly, the symbol grounding problem becomes relevant: we might ground the value in different ways, either by ignoring or considering the decrease in temperature due to increased conduction. Only additional perception and additional knowledge are able to uncover the lack of data quality, because different groundings, which are subject to potentially different aspects of imperfection, are compared.

The question of which information is used in a measure, and how this information is grounded, is related to the question of how the measure is grounded itself. As has been discussed earlier, a quality measure compares the grounding of the original data to another grounding, that is, to the grounding of the measure in the context of the dataset. Accordingly, there exists a duality between the sources of errors in the data, due to a potentially defective grounding, and the quality of the grounding of data quality measures. As a result, the grounding-based ontology of data quality measures that we introduce in this article is complementary to Frank's ontology of data quality [25], because they examine opposing phenomena. The grounding-based ontology examines the grounding of data quality measures, and Frank's ontology of data quality examines the sources of errors when creating data.

The result of the assessment of data quality depends on the compared groundings: the way the data is grounded and the additional grounding by the quality measure. How can data quality be assessed if the assessment is relative to the additionally chosen grounding, albeit data quality itself should be independent of the groundings used during the assessment? If a dataset is assessed several times for its quality, the results of the assessments should ideally coincide, but as assessments are always relative to the chosen groundings, the results of the assessments can differ. For example, the completeness of the OSM road network can be assessed by a comparison to its lineage (*does the length of the road network converge?*), by a comparison to similar areas (*do other residential areas expose a similar structure of houses and roads?*), and by a comparison to (heuristic) principles that are grounded in the environment by our perception (*is the road network connected?*) [48] These different ways of

measuring the completeness of the road network may lead to different judgements about data quality. Even worse, one may question how meaningful such an assessment is: the result of the assessment may depend more on the choice of the compared groundings than on the quality of the data itself. We refer to this phenomenon as the *problem of imperfect data quality assessment*.

The problem of imperfect data quality assessment can be tackled in two ways, either by using groundings that are of good quality with only little imperfection or by using a variety of groundings. In the former case, one may, for example, choose a grounding in a reference dataset whose quality is known to be better than the one of the dataset to be assessed. Data provided by an authoritative source serves in many cases as such a reference. Data that has been less processed and modified and is thus closer to our perception ("raw data") may also be assumed to have a better grounding, because it is closer to the environment. As an example, we might assess the OSM road network by walking around in town and visually perceiving the environment while comparing it to the road network in the data. In the latter case of more than one grounding, it can be hoped that the consideration of a variety of groundings compensates for the imperfection, as statistics does for a set of values that are subject to uncertainty. When the OSM road network is, for example, assessed in respect to its lineage, by comparison to similar areas, by comparison to several (heuristic) principles, and by comparison to several reference datasets, one might easily leave out one of the measures if its result differs considerably from the result of all other principles [48]. The reason behind the measure differing from the other measures might, for example, be a heuristic principle that does not hold under certain circumstances, or a reference dataset of poor quality. The characteristics of measures using heuristic principles differs from the characteristics of measures that use reference data. An understanding of which grounding is used by a measure might reveal these characteristics. In the following section, we discuss an ontology that can be used to classify quality measures according to their groundings.

4   An ontology of quality measures

The assessment of data quality always incorporates the comparison of two or more different groundings of the data and is thus relative to the chosen groundings, as has been argued in the previous section. Hence, it is important to understand which different possibilities exist for grounding data. This understanding renders a more detailed comprehension of data quality assessment and of how to interpret the results of the assessment. In the following, we will systematically explore these different groundings by providing an ontology of quality measures. This ontology classifies measures by the alternative grounding that the data is compared to, and it discusses how these different classes relate.

Data creation and data use define a *perception-action cycle* [54], starting with the perception of and ending with actions in the environment (Figure 1). Data can thus be grounded in two "directions": either by relating to entities that already are grounded in previous perception, for example, to other datasets or to knowledge; or by relating the data to entities that can be grounded by actions in the environment and the perception of the actions' effects on the environment, for example, to processed data. Both directions differ in the way they relate the data to these entities: the first type of a grounding points into the opposite direction of the creation operation (left-facing in Figure 1) while the second type points into the same direction (right-facing).

Figure 1: Groundings of quality measures of *data* (middle row, thick frame). Solid arrows represent the creation of new data from previous data, and grey arrows represent the perception of and the action in the environment. The possibilities to ground quality measures of the data by relating the data to another entity are depicted by dotted arrows. The second row from the top represents *data from the same dataset* as *data*, which may differ in terms of represented aspects, scale, region, the point in time, etc. The last row represents *another dataset*. The depicted classes of the ontology are described in Table 1.

Table 1: Grounding-based ontology of quality measures.

① perception-based grounding	🟩	extrinsic	perception
② data-based grounding	🟨	intrinsic	data from same dataset
	🟨	extrinsic	other dataset
③ grounding in processed data	🟥	intrinsic	processed data from same dataset
	🟥	extrinsic	other processed data
④ grounding in rules/patterns/knowledge	🟨	intrinsic	rules/… gained from data
	🟩	extrinsic	rules/… gained from perception

Consider the example of OSM data. By perceiving the environment, we build a dataset about streets, houses, and many other objects. This data can, for example, be processed to gain route instructions, which can, in turn, be used for navigation in the environment. When the quality of OSM data is to be assessed, it can be compared to authoritative data, which also is the result of a perception of the environment and thus grounded in this perception. The comparison itself, and thus also the assessment, is grounded in this perception as well. The fitness of the data for a certain use can alternatively be assessed by referring to how useful route instructions derived from the data are. That is, how do route instructions derived from OSM data compare to route instructions derived from other data? Both instructions can be compared by determining whether one perceives success after having followed the instructions in the environment. The result of this comparison is influenced by the data and the other dataset, as well as by the algorithms used to derive the route instructions from the datasets. If we can conclude that the influence of the datasets dominates, for example, in the case of using the same algorithm for both datasets, the fitness of the dataset can be assessed. The route instructions themselves can be interpreted as being a dataset as well, because they are the result of having processed the original data.

The groundings of data quality measures are like the grounding of the data itself. Each quality measure compares two different groundings: the original grounding of the data and another alternative grounding. This alternative grounding can uncover data quality issues of the original grounding, and it renders the meaning of the resulting information about the data quality. In fact, the alternative grounding of the data grounds not only the data itself but also their comparison to the original grounding, because it explains the results of the comparison. A data quality measure is thus grounded as well by the alternative grounding. In consequence, an ontology of alternative groundings of data is an ontology of corresponding quality measures, too (Table 1).

Based on Figure 1, the entities in which to ground data, and hence the data quality measures that are grounded in these entities as well, can be assigned the following classes:

Perception-based grounding. Data is created by the perception of the environment, for example, by sensory perceptions, or by different kinds of technical sensors. The quality of the data can be assessed by relating their internal properties to another perception.

Data-based grounding. Data quality can be assessed by a comparison to other data—either data from the same dataset, possibly of another region or another point in time, differing in scale, or containing different thematic information; or to a dataset that has been created independently. A typical example of the latter is authoritative data. The choice of suitable data to compare with is important. If an aspect of data is compared to itself, the comparison becomes meaningless because both compared aspects, which are de facto identical, share the same grounding, and if two very different datasets are compared, it may be hard to relate them in some way.

Grounding in processed data. Instead of the data itself, also specific aspects of processed data can be assessed. Route instructions derived from OSM data can, for example, be compared to route instructions derived from other datasets. If the processed data can be assumed to "continuously" depend on the data, the assessment of the processed data is able to reveal information about the data itself.

Grounding in rules, patterns, or knowledge. Rules, patterns, and knowledge can be gained by analyzing data or our perception. Such derived knowledge is often less prone to error, and quality measures can thus be grounded in the derived knowledge. Statistical information is a typical example.

These classes come, apart from the perception-based grounding, in two guises: **intrinsic and extrinsic groundings**. The grounding used to assess the data is called intrinsic if it incorporates only data that has been created by the same or a very similar perception and extrinsic groundings refer to the opposite. A perception-based grounding is, accordingly, always extrinsic. While the two guises, intrinsic and extrinsic, refer to groundings, they are similar to the concepts of intrinsic/internal and extrinsic/external measures that can be found in literature (e.g., in ISO 19157 about geographical information [39].) As an example, a purely intrinsic grounding refers only to data that was created with a very similar perception, which is, in particular, the case if the measure does not refer to any external dataset and is thus internal according to ISO 19157.

The above classification does not consider any grounding of the data in processed data from another dataset—this would be identical to grounding the data in processed data that itself is compared to processed data of another dataset—nor does it consider any comparison of the data to rules, patterns, or knowledge derived from the processed data—this possibility does not seem to be used very often. Apart from these two exceptions, each choice of a grounding can be uniquely described by one of the preceding classes in combination with the distinction between intrinsic and extrinsic groundings (Table 1).

The proposed ontology does include a taxonomy but also clarifies how the classes relate (Figure 1). An action in the environment starts, for example, according to the ontology, with the perception of the environment, which then leads to data. Such data can be processed in order to finally afford an action in the environment as, for example, route instructions do. Different possibilities of which grounding to refer to in a data quality measure exist. These groundings expose different characteristics, depending on whether they refer to perception, to data, or to processed data. A grounding in rules, patterns, or knowledge also differ considerably from other groundings, because they generalize per definition. These interrelations are depicted in Figure 1.

5   Examples

A grounding-based ontology has been introduced in the preceding section. This ontology applies to data quality measures in general, and it also applies to geographical information and VGI in particular. We provide two examples of quality assessment of VGI with the aim to practically exemplify the ontology. The examples have been chosen such that they illustrate very common ways of measuring the quality of the data, among them intrinsic and extrinsic measures referring to perception, data, and rules.

5.1   Observation of organisms by volunteers

Citizen science projects around the world collect field observations of organisms from volunteers. In the project *ArtenFinder Rheinland-Pfalz*[1], observations of organisms are collected in a dataset with the aim of supporting environmental planning and policy making by government agencies. It is evident that data quality is a vital issue here, because it can have consequences on construction projects, in particular related to species conservation issues. Experts are validating new observations by assessing how plausible a single observation
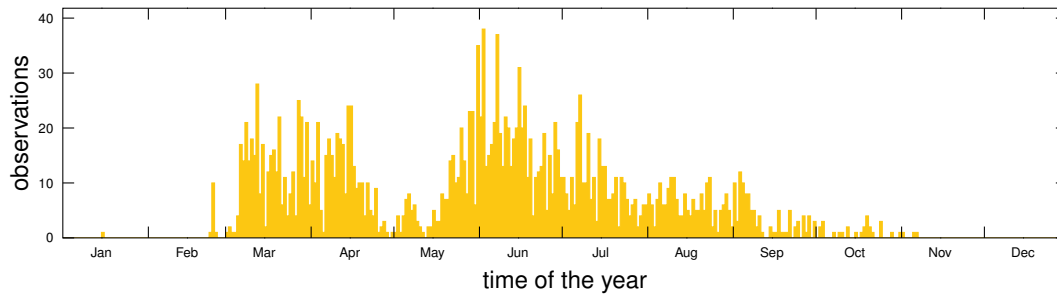
---

[1]http://artenfinder.rlp.de

Figure 2: Histogram of the number of observations of the butterfly species *Aglais urticae* during the year in Rheinland-Pfalz, Germany. The two peaks of observations are the result of hibernating butterflies from the preceding year, which appear in early spring, followed by two new generations from May to November.

is in the context of the existing observations. This process is supported by a tool, which visualizes patterns in the existing validated data [41].

The plausibility check relies on an extrinsic grounding in expert knowledge as well as in the existing, already grounded data. Many species have a variable probability of observation during the year. The corresponding histogram of recorded observations over time (Figure 2) clearly exposes this effect and affords to identify possible seasonal outliers. The interpretation of the histogram (i.e., of existing data) is only meaningful because the recorded observations used to create the histogram are of good quality and have been grounded—we know how observations are made in the environment. The quality assessment of a new observation based on the comparison to this histogram is thus *intrinsic* and *data-based* because it primarily relies on similar observations that have already been stored in the dataset. The quality of the assessment is further improved by combining this grounding with expert knowledge, which is a form of *extrinsic grounding in rules, patterns, or knowledge* (Table 2). While this assessment focuses on temporal aspects, spatial aspects, such as recorded observations of a species in the vicinity of a new observation, can be used in the quality assessment to identify possible ecological outliers.

Table 2: Quality measures in the example of Section 5.1

Measure	Intrinsic/extrinsic	Classification by ontology
expert knowledge	extrinsic	grounding in rules/patterns/knowledge
histogram with the knowledge of patterns in the existing data	intrinsic	data-based grounding

5.2 VGI for supporting crisis management

Crisis management is challenging, and a reaction to a crisis is often needed within hours or days. Information is a major issue for the successful management of a crisis, for example, during natural disasters such as floods, landslides, earthquakes, and tsunamis. Data provided by an authoritative source often serves well for this purpose, but it is increasingly complemented by other data, including OSM data. Like other VGI, OSM relies on the
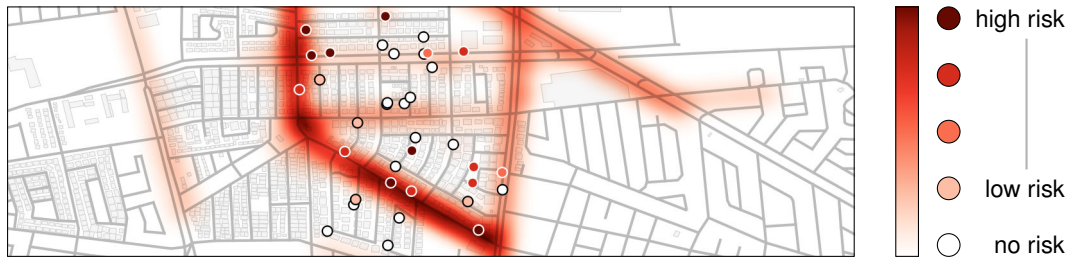
Figure 3: Map depicting the awareness of flood risk in Quilicura, Santiago, Chile. The awareness of the risk in the entire district (heat map) correlates to the awareness of the risk at the interviewed peoples' place of residence (point features).

contribution of volunteers who map data either on the site (local mappers) or by using aerial imagery or information from other sources (remote mappers).

Data quality and the fitness of data are important for crisis management because necessary help and even the life of people may depend on the usefulness of the available information. Mapping a region only with remotely available data sources, for example, aerial imagery, without any need to visit the area of interest, is often advantageous due to the broader availability of mappers, but the lack of local knowledge can lead to misinterpretations or cause mappers to be unable to recognize features of unfamiliar appearance. Remote mappers may, for example, have difficulties with grounding top views of roundhouses or other specific building types unique to a region, if the mapper has no experience with these building types. In contrast, local mappers have common knowledge about these buildings and may map them differently. The quality of remotely created data can thus be assessed by a comparison to locally created data. Such a comparison is an *intrinsic data-based grounding* because local mapping is, in part, similar to remote mapping and the resulting data is stored in the OSM database in both cases. The grounding also exposes *extrinsic* characteristics, because the common knowledge of local and remote mappers are different, and the contexts of the creation process are thus different too [18]. The quality of remotely created data can be improved by conveying local knowledge to remote mappers, for example, by region-specific mapping instructions for typical features of the area of interest [44].

Table 3: Quality measures in the example of Section 5.2

Measure	Intrinsic and extrinsic	Classification by ontology
comparison of remotely and locally created data	intrinsic/extrinsic	data-based grounding
comparison to the perception of flooding at one's own place of residence	extrinsic	perception-based grounding

The creation of data often involves different sensors, for example, GPS devices and aerial images on computers, but human perception is one of the main sources of information. Local mappers use, for example, their visual perception of the surrounding environment and interpret the perception in the context of local knowledge about the area. It is such knowledge that enables the local mapper to, for example, identify a particular object as a roundhouse. Knowledge of a local mapper can also include personal experience of previous

disaster events, which is a valuable source of information within disaster management. A mapper may, for example, identify areas of risks like flooding. Such information can be used to complement authoritative data and data gained by established technical methods. In contrast to technical methods like the detection of flooding by aerial imagery, personal experience provides information about the person's awareness of the risks and is thus clearly subjective. It is of high importance to assess the quality of such information gained from personal experience due to this subjectivity. Consider, for example, data about the awareness of flood risk. Such data can be collected by asking local people to indicate their perception of flood risk at their place of residence, as well as to indicate their perception of flood risk in the entire district. One may assume that the perception of flood risk is more accurate at a person's own place of residence because the person's own property is affected. Given that this assumption holds, the quality of the data gained by the perception of flood risk in the district can be assessed by a comparison to the data at the person's place of residence [45], which is an example of a *perception-based grounding* (Figure 3; Table 3).

6   Evaluation of the ontology

We have introduced an ontology of quality measures in the preceding sections, and we have discussed how the ontology applies to practical examples. This section is dedicated to an evaluation of the ontology. There exists a multiplicity of quality measures for OSM data, yet a detailed and meaningful classification in respect to the grounding of the data is missing in literature. We demonstrate that the ontology indeed yields a meaningful classification of quality measures in the case of OSM data (Table 4), by providing a good overview over the different types of quality measures and by identifying missing types of quality measures.

The creation of Table 4 incorporates a selection of quality measures for OSM data, a selection which influences the result of the evaluation of the ontology. In order to make this selection comprehensible and more reproducible, we focus only on these measures that are well documented in the scientific literature. Measures that are only used by the OSM community are, for example, not considered. More specifically, we have only included publications and measures that meet the following criteria:

- The publication has either been listed in Google Scholar before June 1, 2017, or the publication is provided in the references of an article that meets the other criteria and has been listed in Google Scholar before June 1, 2017.

- The publication is either published as a journal article, as a conference paper or as a book chapter.

- We have full-text access to the publication, which is why we excluded theses in general.

- The publication clearly refers to "OpenStreetMap" or "OSM" in combination with "data quality," "fitness for purpose," or similar terms either in the title, the abstract, or among the keywords. In addition, review articles were also included if they refer to "volunteered geographic information" or "VGI" instead of "OpenStreetMap" or "OSM."

- If a measure has been published by the same author in very similar publications, we have only included it once.

Table 4: Examples of quality measures for OpenStreetMap data.

perception-based grounding	extrinsic	comparison to the perception of experts [18]
		comparison to a survey of roads by photogrammetric methods [11]
data-based grounding	intrinsic	comparison of data at different points in time to assess consistency [36]
		comparison to data in the neighborhood to assess thematic quality [63]
	extrinsic	comparison to other data, among them official survey data, expert data, and proprietary data, to assess completeness, spatial accuracy, and thematic accuracy [3, 10, 13, 19, 20, 23, 28, 33, 37, 38, 40, 46, 51, 61]
grounding in processed data	intrinsic	–
	extrinsic	assessing completeness of the road network by comparison of routing results to Google Maps [57]
grounding in rules/patterns/ knowledge	intrinsic	logical consistency [28]
		logical consistency for landuse [5]
		plausibility [1]
		saturation principle as an indicator for completeness: generally [55], and of roads [5]
		absolute number of points of interests (POIs) as an indicator for the completeness of the data [5]
		conceptual quality [2, 4]
		geometric and thematic lineage [2, 61]
		thematic completeness for house numbers [5]
		spatial accuracy of roads: number of contributors as an indicator for data quality (Linus' law) [34]
		patterns in the mapping behavior relate to data quality [6, 32, 43, 52]
		patterns in the tagging behavior relate to data quality [50, 57]
	extrinsic	plausibility in regards to spatial relations that are derived from our conceptualization of the environment [61]
		providing tagging guidance by OSM editor [15, 63]

- Only measures that are explained in detail in one of these publications have been included. The measures were, however, not filtered by its quality, because the quality of a data quality measure is hard to formalize and measure.

Based on the above criteria, the process of selecting these publications and data quality measures contained within can be described as a *systematic literature review*.

How should some quality measure practically be classified? Each measure compares one or more alternative groundings to the "given" interpretation[2] of the data. It is this alternative grounding which is classified according to the ontology, but a classification often cannot be made without a doubt. The assessment of the number of contributors, for example, can be seen as an indicator for data quality. This measure could be classified as a comparison to an *intrinsic* grounding in a rule (the rule can be derived by the analysis of the OSM data itself) or it could be classified as a comparison to an *extrinsic* grounding in a rule (the rule can be derived by considerations independent of the actual OSM data). Similar issues seem, in fact, to arise for most of the groundings in rules, patterns, or knowledge. Another ambiguous example is the assessment for logical consistency. It could be classified as a comparison to a *data-based grounding* or to a *grounding in a rule*, because a comparison to the entire dataset as well as to some formal rules of logic are incorporated. Also this issue can be seen as being very typical, because almost every grounding incorporates some basic logic. Conceptual quality of a dataset can be measured by a grounding in rules, patterns, or knowledge. Even if the quality can be measured *intrinsically*, as different examples demonstrate [2, 4, 49], the assessment also incorporates our conceptualization, which refers to the environment and is thus of *extrinsic* nature. Finally, ambiguity also arises when a measure refers to several data sources.

One can face ambiguity in several ways when classifying quality measures, but as ambiguity is an inherent characteristic of data quality measures, it cannot be avoided. This applies not only to our ontology but also to other ones, for example, Frank's ontology of data quality [25] and ISO 19157 [39]. In fact, measures often examine combinations of different aspects of data quality. As an example, the comparison to a reference dataset often reveals positional and thematic accuracy as well as completeness, etc. In the classification of the quality measures in Table 4, we have assigned only one class to each measure, which makes the table less complex. In cases where several classes apply to a measure, the most suitable one has been chosen. Some articles occur more than once in Table 4 because they describe more than one measure. The issue of choosing one class remains though often a gradual one, as can be seen in the following example: Canavosio-Zuzelski et al. [11] assess the road network represented within OSM data by a comparison to aerial stereo imagery. This comparison, regarded as a quality measure, can be classified as an extrinsic data-based grounding (data is recorded), and also as a perception-based grounding (it refers to another method of perception). The measure fits into both classes, but relying on a different way of perception seems to be more important because this increases the chance to detect errors.

The classification of the quality measures in Table 4 evaluates several aspects, the quality and the usefulness of the ontology as well as the evaluation of the selection of measures. Several qualities of ontologies and taxonomies have been discussed by Fernández et al. [21] and Burton-Jones et al. [9]. Among these qualities are a meaningful scope and a uniform granularity. If suitable examples exist for each class, the scope of the ontology fits to the scope of the examined selection of quality measures. If the number of examples for each class are about equal, the granularity of the ontology fits to the granularity of the examined selection. These two aspects demonstrate that the usefulness of the ontology and the evaluation of the selection of measures relate. In the case of a class to which very few or very many

---

[2]A quality measure can only assess the quality of an interpretation of the data, not of the data itself, because data only becomes meaningful when being interpreted. While there is no "given" interpretation of data in general, we assume a fixed natural interpretation of the data in the context of this evaluation.
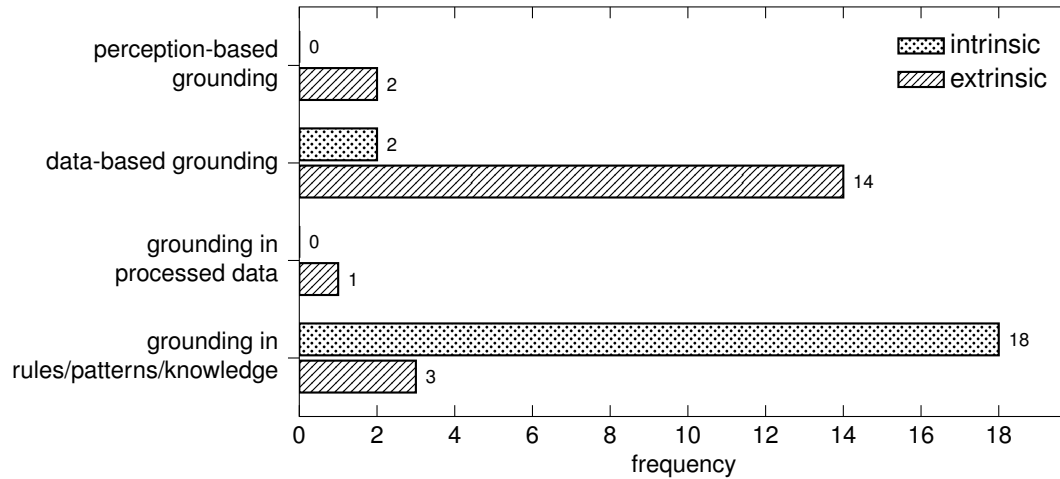
Figure 4: Frequency distribution of the measures in Table 4, classified according to the grounding-based ontology.

measures are assigned, the meaningfulness of the class needs to be assessed independently. The frequency distribution of the measures is depicted in Figure 4.

Two of the main classes contain considerably more measures than the other two classes in Table 4. There exist sixteen data-based groundings and twenty-one groundings in rules/patterns/knowledge, whereas only two perception-based groundings and one grounding in processed data are listed. As most of the measures classified as groundings in rules/patterns/knowledge are intrinsic (eighteen measures), the majority of measures rely on data. Measures that rely on the perception of the environment or on processed data are in the minority. This unequal distribution could be interpreted as a weakness of the ontology but can easily be explained in terms of the selection. The class of intrinsic groundings in rules, patterns, or knowledge, for example, is the one that most measures are assigned to. One of the reasons might be that intrinsic measures are very efficient because they only rely on the data itself. Research focuses broadly on such intrinsic methods, and the intrinsic measures that are grounded in rules, patterns, or knowledge are, accordingly, very diverse (Table 4). This is in contrast to a comparison to existing data (often reference datasets), because such a comparison needs much less effort than a new perception of the environment, or than a comparison to processed data. The class of extrinsic data-based groundings is, accordingly, the one that the second-most measures are assigned to. This is not by chance either. Many articles introduce intrinsic measures, which are then usually assessed by an extrinsic data-based measure, because a comparison to reference data seems to suggest itself for this purpose. Many of these data-based measures are, accordingly, very similar or even identical.

Only a minority of measures rely on the perception of the environment, or on processed data, as previously discussed. The perception and the processed data are closer to the environment (Figure 1), which can be seen as follows. First, we perceive the environment, which results in raw data. Then, we use the raw data to build a dataset, for example, the dataset that is to be examined. This dataset, or a comparable one, can be processed again, usually with the purpose to allow for an action in the environment. As an example, we

perceive the environment through using aerial images and add new information to the OSM dataset. Routing algorithms can utilize OSM data in order to compute routes, which is an example of processed data. These route instructions can, in turn, afford the navigation to the desired place in the environment. Aerial images and routing instructions are thus closer to the environment than OSM data in the sense that they are directly related to the environment either by perception or by actions. It has been argued before that the easier accessibility of OSM data compared to the perception and processed data might be a reason for the dominance of data-based measures. Despite this dominance, examples of measures exist for the other classes, apart from measures using an intrinsic grounding in processed data.

Measures that rely on perception or on processed data are underrepresented in the literature, despite being closer to the environment. Processed data like route instructions usually reflects a certain purpose, because they are mostly created with the intention to give rise to an action in the environment. A measure that is grounded in processed data can thus be expected to optimally afford the assessment of the fitness for purpose. A comparison of route instructions generated by one algorithm using different datasets might, for example, offer valuable clues about how fit the datasets are for navigation tasks. In contrast, measures relying on perception contain maximal information about the environment and can be expected to depend less on a purpose. Such measures are thus useful for assessing data quality independent of a certain purpose.

The evaluation shows two major findings. First, measures exist in most classes, and very different measures can easily be distinguished by their classification, which demonstrates that the ontology is of a suitable granularity. The class of measures that are intrinsically grounded in rules, patterns, or knowledge contains many different measures though, which can be further classified by ISO 19157 [39] according to which aspects they assess. Secondly, the ontology provides a view on ongoing research areas and reveals potential future research areas: the examination of existing and development of new measures that are grounded in the perception of the environment; in rules, patterns, or knowledge that has been derived by our perception of the environment; or in processed data.

7   Conclusion

We examined *why* data quality measures work; that is, what information measures use to assess the quality of data. The used information provides an alternative grounding of the data, which potentially refers to the environment in a different way than the original grounding. This, in turn, sheds light on the quality of the data. Different types of such alternative groundings have been identified and set into context, which resulted in a grounding-based ontology. The ontology has been critically evaluated in the context of OpenStreetMap data.

The processes of perceiving the environment, creating datasets, and processing these in order to afford an action in the environment have been discussed in the context of how data quality measures can be grounded. This chain of processes could be examined with a much finer granularity, as is done, among others, in Frank's data quality ontology. These smaller "steps" cause different types of error, which can be assessed in different ways and using different groundings. Future work may discuss the characteristics of these smaller steps and argue which classes of measures are, according to the grounding-based ontology, most suitable for assessing each of these steps with respect to data quality. These smaller steps

may, in particular, involve very essential concepts of geographical information like scale and level of detail, granularity, uncertainty, etc. How do these concepts relate to different groundings, and how can quality be assessed in respect to these concepts?

Intrinsic data quality measures only refer to data that has been created in a similar way as the data to assess. While intrinsic measures seem to lack an alternative grounding because they do not refer to additional information, they cannot exist without such a grounding—data quality cannot be assessed without comparing different groundings. Future research may examine intrinsic measures in more detail to gain insights about which presumptions are implicitly made and which alternative groundings are implicitly used.

The evaluation revealed that in literature only very few data quality measures that rely on perception or on processed data have been discussed. In particular, there exists a lack of data quality measures that compare processed data despite the common aim to assess the fitness for purpose in the context of applications. In fact, quality measures that are grounded in processed data can be expected to deliver insightful results and strongly contribute to the assessment of fitness for purpose. Future research may fill this gap and new measures may be developed.

The grounding-based ontology can improve our understanding of how different data quality measures correlate, and how they can mutually complement each other. In particular, a better understanding of how different measures contribute to the assessment of a specific data quality dimension can be gained if the measures are classified according to the grounding-based ontology in order to compare the involved groundings. How can shortcomings of single measures be compensated by other measures? How can the combination of two or more measures lead to stronger results of the quality assessment? What can be followed from two similar measures resulting in contradicting findings? How can the reliability of two measures be compared? Answers to such questions may lead to the development of more sophisticated algorithms to automate quality assessment. In addition, the questions may lead to a theory of how to describe and measure the quality of data quality measures.

There exists a duality between the grounding-based ontology introduced in this article and Frank's data quality ontology, as has been argued. Furthermore, ISO 19157 has been shown to complement the grounding-based ontology. Different ontologies about data quality, fitness for purpose, and data quality measures have been discussed, but a more thorough understanding of how these ontologies relate is still needed. Can these ontologies be merged into a single ontology? Do these ontologies overlap? Can the duality between the grounding-based ontology and Frank's data quality ontology be formalized? These questions may contribute to the comprehension of the issue of data quality and finally lead to more elaborate definitions of data quality and fitness for purpose.

Author contributions

Acknowledgments

References

[1] ALI, A. L., SCHMID, F., AL-SALMAN, R., AND KAUPPINEN, T. Ambiguity and plausibility: managing classification quality in volunteered geographic information. *Proc. of the 22nd ACM International Conference on Advances in Geographic Information Systems, SIGSPATIAL* (2014), 143–152. doi:10.1145/2666310.2666392.

[2] ANTONIOU, V., TOUYA, G., AND RAIMOND, A.-M. Quality analysis of the Parisian OSM toponyms evolution. In *European Handbook of Crowdsourced Geographic Information*, C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, Eds. Ubiquity Press, London, 2016, pp. 97–112. doi:10.5334/bax.h.

[3] ARSANJANI, J. J., MOONEY, P., ZIPF, A., AND SCHAUSS, A. Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets. In *OpenStreetMap in GIScience. Experiences, research, and applications*, J. J. Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Springer, Heidelberg, 2015, pp. 37–58. doi:10.1007/978-3-319-14280-7_3.

[4] BALLATORE, A., AND ZIPF, A. A conceptual quality framework for volunteered geographic information. *Proc. of the 12th Conference on Spatial Information Theory, COSIT* (2015), 89–107. doi:10.1007/978-3-319-23374-1_5.

[5] BARRON, C., NEIS, P., AND ZIPF, A. A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Transactions in GIS 18*, 6 (2014), 877–895. doi:10.1111/tgis.12073.

[6] BÉGIN, D., DEVILLERS, R., AND ROCHE, S. Assessing volunteered geographic information (VGI) quality based on contributors' mapping behaviours. *Proc. of the 8th International Symposium on Spatial Data Quality* (2013), 149–154. doi:10.5194/isprsarchives-XL-2-W1-149-2013.

[7] BILJECKI, F., HEUVELINK, G. B. M., LEDOUX, H., AND STOTER, J. The effect of acquisition error and level of detail on the accuracy of spatial analyses. *Cartography and Geographic Information Science* (2017). doi:10.1080/15230406.2017.1279986.

[8] BORDOGNA, G., CARRARA, P., CRISCUOLO, L., PEPE, M., AND RAMPINI, A. On predicting and improving the quality of volunteer geographic information projects. *International Journal of Digital Earth 9*, 2 (2014), 134–155. doi:10.1080/17538947.2014.976774.

[9] BURTON-JONES, A., STOREY, V. C., SUGUMARAN, V., AND AHLUWALIA, P. A semiotic metrics suite for assessing the quality of ontologies. *Data and Knowledge Engineering 55*, 1 (2005), 84–102. doi:10.1016/j.datak.2004.11.010.

[10] CAMBOIM, S. P., BRAVO, J. V. M., AND SLUTER, C. R. An investigation into the completeness of, and the updates to, OpenStreetMap data in a heterogeneous area in Brazil. *ISPRS International Journal of Geo-Information 4* (2015), 1366–1388. doi:10.3390/ijgi4031366.

[11] CANAVOSIO-ZUZELSKI, R., AGOURIS, P., AND DOUCETTE, P. A photogrammetric approach for assessing positional accuracy of OpenStreetMap roads. *ISPRS International Journal of Geo-Information 2*, 2 (2003), 276–301. doi:10.3390/ijgi2020276.

[12] CHRISMAN, N. R. The role of quality information in the long-term functioning of a geographic information system. *Cartographica 21*, 2 (1984), 79–87.

[13] CIPELUCH, B., JACOB, R., WINSTANLEY, A. C., AND MOONEY, P. Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. *Proc. of the 9th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, ACCURACY* (2010), 337–340.

[14] COUCLELIS, H. The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS 7*, 2 (2003), 165–175. doi:10.1111/1467-9671.00138.

[15] DAVIDOVIC, N., MOONEY, P., STOIMENOV, L., AND MINGHINI, M. Tagging in volunteered geographic information: an analysis of tagging practices for cities and urban regions in OpenStreetMap. *ISPRS International Journal of Geo-Information 5*, 12 (2016). doi:10.3390/ijgi5120232.

[16] DE ALBUQUERQUE, J. P., HERFORT, B., BRENNING, A., AND ZIPF, A. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science 29*, 4 (2015), 667–689. doi:10.1080/13658816.2014.996567.

[17] DEVILLERS, R., BÉDARD, Y., AND JEANSOULIN, R. Multidimensional management of geospatial data quality information for its dynamic use within GIS. *Photogrammetric Engineering and Remote Sensing 71*, 2 (2005), 205–215. doi:10.14358/PERS.71.2.205.

[18] ECKLE, M., AND DE ALBUQUERQUE, J. P. Quality assessment of remote mapping in OpenStreetMap for disaster management purposes. *Proc. of the 12th International Conference on Information Systems for Crisis Response and Management, ISCRAM* (2015).

[19] FAN, H., YANG, B., ZIPF, A., AND ROUSELL, A. A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. *International Journal of Geographical Information Science 30*, 4 (2016), 748–764. doi:10.1080/13658816.2015.1100732.

[20] FAN, H., ZIPF, A., FU, Q., AND NEIS, P. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science 28*, 4 (2014), 700–719. doi:10.1080/13658816.2013.867495.

[21] FERNÁNDEZ, M., OVERBEEKE, C., SABOU, M., AND MOTTA, E. What makes a good ontology? A case-study in fine-grained knowledge reuse. *Proc. of the 4th Asian Conference on The Semantic Web, ASWC* (2009), 61–75. doi:10.1007/978-3-642-10871-6_5.

[22] FONTE, C. C., ANTONIOU, V., BASTIN, L., ESTIMA, J., ARSANJANI, J. J., BAYAS, J.-C. L., SEE, L., AND VATSEVA, R. Assessing VGI data quality. In *Mapping and the citizen sensor*, G. M. Foody, L. See, S. Fritz, P. Mooney, A.-M. Olteanu-Raimond, C. C. Fonte, and V. Antoniou, Eds. Ubiquity Press, London, 2017, pp. 137–163. doi:10.5334/bbf.g.

[23] FORGHANI, M., AND DELAVAR, M. R. A quality study of the OpenStreetMap dataset for Tehran. *ISPRS International Journal of Geo-Information 3*, 2 (2014), 750–763. doi:10.3390/ijgi3020750.

[24] FRANK, A. U. Metamodels for data quality description. In *Data quality in geographic information. From error to uncertainty*, R. Jeansoulin and M. F. Goodchild, Eds. HermÃĺs, Paris, 1998, pp. 15–29.

[25] FRANK, A. U. Data quality ontology: an ontology for imperfect knowledge. *Proc. of the 8th Conference on Spatial Information Theory, COSIT* (2007), 406–420. doi:10.1007/978-3-540-74788-8_25.

[26] FRANK, A. U. Why is scale an effective descriptor for data quality? The physical and ontological rationale for imprecision and level of detail. In *Research trends in geographic information science*, G. Navratil, Ed. Springer, Heidelberg, 2009, pp. 39–61. doi:10.1007/978-3-540-88244-2_4.

[27] GIBSON, J. J. The theory of affordances. In *Perceiving, acting, and knowing*, R. Shaw and J. Bransford, Eds. Lawrence Erlbaum Associates, Hillsdale, NJ, 1977, pp. 127–143.

[28] GIRRES, J.-F., AND TOUYA, G. Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS 14*, 4 (2010), 435–459. doi:10.1111/j.1467-9671.2010.01203.x.

[29] GOODCHILD, M. F. Geographical information science. *International Journal of Geographical Information Systems 6*, 1 (1992), 31–45. doi:10.1080/02693799208901893.

[30] GOODCHILD, M. F. Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science 1* (2010), 3–20. doi:10.5311/JOSIS.2010.1.2.

[31] GOODCHILD, M. F., AND LI, L. Assuring the quality of volunteered geographic information. *Spatial Statistics 1* (2012), 110–120. doi:10.1016/j.spasta.2012.03.002.

[32] GRÖCHENIG, S., BRUNAUER, R., AND REHRL, K. Estimating completeness of VGI datasets by analyzing community activity over time periods. *Proc. of the 17th AGILE Conference on Geographic Information Science* (2014), 3–18. doi:10.1007/978-3-319-03611-3_1.

[33] HAKLAY, M. How good is volunteered goegraphical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B 37*, 4 (2010), 682–703. doi:10.1068/b35097.

[34] HAKLAY, M., BASIOUKA, S., ANTONIOU, V., AND ATHER, A. How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. *The Cartographic Journal 47*, 4 (2010), 315–322. doi:10.1179/000870410X12911304958827.

[35] HARNAD, S. The symbol grounding problem. *Physica D 42*, 1–3 (1990), 335–346. doi:10.1016/0167-2789(90)90087-6.

[36] HASHEMI, P., AND ABBASPOUR, R. A. Assessment of logical consistency in Open-StreetMap based on the spatial similarity concept. In *OpenStreetMap in GIScience. Experiences, research, and applications*, J. J. Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Springer, Heidelberg, 2015, pp. 19–36. doi:10.1007/978-3-319-14280-7_2.

[37] HECHT, R., KUNZE, C., AND HAHMANN, S. Measuring completeness of building footprints in OpenStreetMap over space and time. *ISPRS International Journal of Geo-Information 2*, 4 (2013), 1066–1091. doi:10.3390/ijgi2041066.

[38] HELBICH, M., AMELUNXEN, C., NEIS, P., AND ZIPF, A. Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary geodata. *Proc. of the GI_Forum* (2012), 24–33.

[39] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 19157:2013. Geographic information. Data quality, 2013.

[40] JACKSON, S. P., MULLEN, W., AGOURIS, P., CROOKS, A., CROITORU, A., AND STEFANIDIS, A. Assessing completeness and spatial error of features in volunteered geographic information. *ISPRS International Journal of Geo-Information 2*, 2 (2013), 507–530. doi:10.3390/ijgi2020507.

[41] JACOBS, C., AND SCHOTTHÖFER, A. Citizen-Science-Daten zur Biodiversität: Methoden zur Unterstützung der Qualitätssicherung. *Proc. of the Symposium und Expo Angewandte Geoinformation, AGIT* (2015).

[42] JACOBS, C., AND ZIPF, A. Completeness of citizen science biodiversity data from a volunteered geographic information perspective. *Geo-spatial Information Science 20*, 1 (2017), 3–13. doi:10.1080/10095020.2017.1288424.

[43] KESSLER, C., AND DE GROOT, R. T. A. Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. *Proc. of the 16th AGILE Conference on Geographic Information Science* (2013), 21–37. doi:10.1007/978-3-319-00615-4_2.

[44] KLONNER, C., ECKLE, M., USÓN, T., AND HÖFLE, B. Quality improvement of remotely volunteered geographic information via country-specific mapping instructions. *Proc. of the 14th International Conference on Information Systems for Crisis Response and Management, ISCRAM* (2017).

[45] KLONNER, C., MARX, S., USÓN, T., AND HÖFLE, B. Risk awareness maps of urban flooding via OSM field papers – case study Santiago de Chile. *Proc. of the 13th International Conference on Information Systems for Crisis Response and Management, ISCRAM* (2016).

[46] KOUKOLETSOS, T., HAKLAY, M., AND ELLUL, C. Assessing data completeness of VGI through an automated matching procedure for linear data. *Transactions in GIS 16*, 4 (2012), 477–498. doi:10.1111/j.1467-9671.2012.01304.x.

[47] MOCNIK, F.-B. *A scale-invariant spatial graph model*. PhD thesis, Vienna University of Technology, 2015.

[48] MOCNIK, F.-B., ZIPF, A., AND FAN, H. The inevitability of calibration in VGI quality assessment. *Proc. of the 4th Workshop on Volunteered Geographic Information: Integration, Analysis, and Applications, VGI-Analytics* (2017).

[49] MOCNIK, F.-B., ZIPF, A., AND RAIFER, M. The OpenStreetMap folksonomy and its evolution. *Geo-spatial Information Science 20*, 3 (2017), 219–230. doi:10.1080/10095020.2017.1368193.

[50] MOONEY, P., AND CORCORAN, P. The annotation process in OpenStreetMap. *Transactions in GIS 16*, 4 (2012), 561–579. doi:10.1111/j.1467-9671.2012.01306.x.

[51] NEIS, P., ZIELSTRA, D., AND ZIPF, A. The street network evolution of crowd-sourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet 4* (2012), 1–21. doi:10.3390/fi4010001.

[52] NEIS, P., ZIELSTRA, D., AND ZIPF, A. Comparison of volunteered geographic information data contributions and community development for selected world regions. *Future Internet 5*, 2 (2013), 282–300. doi:10.3390/fi5020282.

[53] OGDEN, C. K., AND RICHARDS, I. A. *The meaning of meaning*. Harcourt, Brace and World, New York, 1923.

[54] ORTMANN, J. *Semantic integration of human and technical observations*. PhD thesis, University of Münster, 2014.

[55] REHRL, K., AND GRÖCHENIG, S. A framework for data-centric analysis of mapping activity in the context of volunteered geographic information. *ISPRS International Journal of Geo-Information 5*, 37 (2016). doi:10.3390/ijgi5030037.

[56] SCHEIDER, S. *Grounding geographic information in perceptual operations*. IOS Press, Amsterdam, 2012.

[57] SEHRA, S. S., SINGH, J., AND RAI, H. S. Assessing OpenStreetMap data using intrinsic quality indicators: an extension to the QGIS processing toolbox. *Future Internet 9*, 2 (2017). doi:10.3390/fi9020015.

[58] SENARATNE, H., MOBASHERI, A., ALI, A. L., CAPINERI, C., AND HAKLAY, M. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science 31*, 1 (2017), 139–167. doi:10.1080/13658816.2016.1189556.

[59] SHI, W., FISHER, P. F., AND GOODCHILD, M. F., Eds. *Spatial data quality*. Taylor and Francis, London, 2002.

[60] SWISS CONFEDERATION, AIRCRAFT ACCIDENT INVESTIGATION BUREAU. Schlussbericht Nr. 1793 des Büros für Flugunfalluntersuchungen.

[61] TOUYA, G., ANTONIOU, V., OLTEANU-RAIMOND, A.-M., AND VAN DAMME, M.-D. Assessing crowdsourced POI quality: combining methods based on reference data history, and spatial relations. *ISPRS International Journal of Geo-Information 6*, 80 (2017). doi:10.3390/ijgi6030080.

[62] VAN OORT, P. *Spatial data quality: from description to application*. Netherlands Geodetic Commission, Delft, 2005.

[63] VANDECASTEELE, A., AND DEVILLERS, R. Improving volunteered geographic information quality using a tag recommender system: the case of OpenStreetMap. In *OpenStreetMap in GIScience. Experiences, research, and applications*, J. J. Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Springer, Heidelberg, 2015, pp. 59–80. doi:10.1007/978-3-319-14280-7_4.

[64] WAND, Y., AND WANG, R. Y. Anchoring data quality dimensions in ontological foundations. *Communications of the ACM 39*, 11 (1996), 86–95. doi:10.1145/240455.240479.