A Framework of Quality Assessment Methods for Crowdsourced Geographic Information: a Systematic Literature Review

Lívia Castro Degrossi
Department of Computer Systems
University of São Paulo
São Carlos, Brazil
degrossi@usp.br

João Porto de Albuquerque
Centre for Interdisciplinary Methodologies
University of Warwick, Coventry, UK
GIScience Research Group
Heidelberg University, Germany
j.porto@warwick.ac.uk

Roberto dos Santos Rocha
Department of Computer Systems
University of São Paulo
São Carlos, Brazil
rsrocha@usp.br

Alexander Zipf
GIScience Research Group
Heidelberg University
Heidelberg, Germany
zipf@uni-heidelberg.de

ABSTRACT

Crowdsourced Geographic Information (CGI) has emerged as a potential source of geographic information in different application domains. Despite the advantages associated with it, this information lacks quality assurance, since it is provided by different people. Therefore, several authors have started investigating different methods to assess the quality of CGI. Some of the existing methods have been summarized in different classification scheme. However, there is not an overview of the methods employed to assess the quality of CGI in the absence of authoritative data. On the basis of a systematic literature review, we found 13 methods that can be employed to this end.

Keywords

Volunteered Geographic Information, VGI, Crowdsourced Geographic Information, Quality Assessment, Systematic Literature Review.

INTRODUCTION

The use of Crowdsourced Geographic Information (CGI) has grown in the past few years owing to a number of key features, e.g. it is free, up-to-date and provided by several volunteers. CGI is an umbrella term that encompasses both "active/conscious" and "passive/unconscious" georeferenced information (See et al., 2016). This term has been used as an equivalent of Volunteered Geographic Information (VGI) (Goodchild, 2007) since some researchers have questioned the use of the term "volunteered" to refer to information that is collected without the will or conscious knowledge of the provider (Harvey, 2013). VGI is georeferenced information that is produced by volunteers using appropriate tools like Web portals and mobile devices (Goodchild, 2007). This type of information can be obtained through three collaborative activities: (*i*) social media, (*ii*) crowd sensing and (*iii*) collaborative mapping (Albuquerque, Herfort, Eckle, & Zipf, 2016).

When making use of CGI, ensuring the quality of the information is a challenging question. The information that is supplied by volunteers does not have to comply with any quality standards, and there is no control of the creation process. In addition, most volunteers do not have any formal training or cartographic skills. It is thus becoming increasingly important to assess the quality of CGI before it is used. Quality assessment is an important step to understanding if the information is fit-for-purpose with regard to the way it will be used (Ballatore & Zipf, 2015).

Several researchers have started investigating different approaches to assess the quality of CGI. Currently, there are a large number of methods to accomplish this task (e.g. Foody et al., 2013; Girres & Touya, 2010). These methods differ with regard to the type of information evaluated, and reference data types, among other factors. Owing to the large number of existing methods, selecting one is not a trivial task.

In an attempt to summarize these methods, several researchers have reviewed and categorized them in the literature (Bordogna et al., 2016; Mirbabaie et al., 2016; Senaratne et al., 2017; Wiggins et al., 2011). They have sought to answer questions such as the following: What are the methods used to assess the quality of CGI in Citizen Science projects? What are the methods to assess the quality of map-, image-, and text-based CGI? However, a question still remains. What are the methods used to assess the quality of CGI in the absence of authoritative data? This is important since authoritative data may not be available or may be out-of-date. Authoritative data are ground-truth data that are provided by official and trustable sources.

In this work, we address this question by carrying out a Systematic Literature Review (SLR) to discover the existing methods to assess the quality of CGI when authoritative data is not available. This SLR aims at providing an overview of the characteristics of the existing methods for researchers and developers of crowdsourcing-based platforms. Additionally, this work can be used to identify where further investigation is still needed.

The remainder of this article is structured as follows. In Section 2, we examine the quality of CGI. In Section 3, there is an overview of related works. In Section 4, the methodology employed for carrying out the SLR. Following this, each method is outlined in detail in Section 5. In Section 6, we discuss and summarize our findings and the limitations of our SLR, and make suggestions for future research.

QUALITY OF CGI

When dealing with crowdsourced geographic information, concerns arise regarding its quality and value as an information source. The quality of CGI largely depends on different factors such as the characteristics of the volunteer, the type of information, and the way in which the information is produced (Bordogna et al., 2016).

CGI is provided by a wide range of sources (e.g. volunteers), who have different levels of expertise and come from different backgrounds. There are several reasons for supplying information and heterogeneous methods for interpreting and communicating it. Volunteers are responsible for providing types of information that are characterized by a lack of structure, since it is collected by several crowdsourcing platforms with heterogeneous media formats and interface options. Furthermore, this kind of information is produced without a standard process. A combination of these factors leads to heterogeneous quality, which can affect the usability of the crowdsourced information (Bishr & Kuhn, 2013). However, the effects of its quality depends on how the information is used (Goodchild & Glennon, 2010), since the quality of the information is determined by the context in which it is applied (Bordogna et al., 2016). Depending on the context, the quality of CGI may be a serious factor (Goodchild & Glennon, 2010). Hence, a knowledge of the degree of quality allows people to use the information with confidence or may warn them of potential risks (Jilani et al., 2014).

The quality of CGI can be measured by means of distinct elements. A "quality element" is a component that describes an aspect of the quality of geographic information (ISO, 2013). The International Organization for Standardization (ISO) defines a set of quality elements that can be used to measure the quality of geographic information, i.e. completeness, positional accuracy, thematic accuracy, logical consistency, temporal quality and usability.

These quality elements are used to measure the quality of CGI. However, this type of information has specific features which makes conducting a quality assessment different from the case of traditional geographic data (Mohammadi & Malek, 2015). Hence, researchers have added new elements to measure CGI quality, such as trust, or made new definitions for existing quality elements, which can be measured by different methods (see the description in Section 5). They are characterized in line with (i) the approach adopted (i.e., geographic, crowdsourcing, and social); (ii) the type of reference data, i.e. intrinsic or extrinsic; and (iii) the temporality of the method (ex ante vs. ex post). Each category is described in the following sections.

Approaches for quality assessment of CGI

Goodchild & Li (2012) proposed three approaches to assess CGI quality: (i) crowdsourcing, (ii) social and (iii) geographic. The crowdsourcing approach is based on the ability of a group to detect and correct possible mistakes made by an individual. In OpenStreetMap, for example, it is possible to edit and correct erroneous geographic features provided by other people.

In the social approach, people at a higher level in the hierarchy act as gatekeepers of crowdsourcing platforms. Thus, a group of people can maintain the integrity of the platform, prevent vandalism and copyright protected material, and avoid abusive content. In the Flood Citizen Observatory (Degrossi et al., 2014), for instance, the platform administrator acts as a gatekeeper, by assessing the veracity of CGI and classifying it as checked or unchecked.

Finally, in the geographic approach, CGI is compared with a geographic dataset. This approach is based on the First Law of Geography, where "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Thus, a geographic information is more related to the geographic context of that area consistent with it historical facts. Albuquerque et al. (2014) showed, for instance, that when the overall number of flood-related tweets are compared, "there is perhaps a tendency for 'relevant' on-topic tweets to be closer to flood-affected catchments".

Reference Data

Depending on which reference dataset is used, methods for quality assessment can be classified as either extrinsic or intrinsic. Extrinsic methods use external knowledge to measure the quality of CGI. Although authoritative data are commonly used as external knowledge, their use can be constrained by financial costs, licensing restrictions (Mooney et al., 2010), and currency (Goodchild & Li, 2012).

Intrinsic methods do not rely on external knowledge for assessing the quality of CGI. Despite this, these methods are employed to analyze historical metadata as a means of inferring the inherent quality of the data. Thus, it is possible to evaluate the quality of CGI regardless of whether a reference dataset is available or not. However, in some cases, intrinsic methods do not allow absolute statements to be made about CGI quality. Thus, they can only be used for making rough estimates of the possible data quality (Barron et al., 2014).

Temporality

The assessment of the quality of CGI can be carried out in light of two temporalities: (i) ex ante and (ii) ex post (Bordogna et al., 2016). These differ with regard to the time when the assessment is carried out compared with the creation time of CGI. The *Ex ante* strategy is employed before a CGI is created and seeks to avoid the creation of low-quality CGI (Bordogna et al., 2016). As well as offering mechanisms for controlling data creation, these methods also provide volunteers with resources for guiding the way information is produced. In contrast, the *Ex post* strategy is employed after a CGI item has been created. This strategy aims at removing and improving CGI quality. This involves first checking the quality of CGI and, later, filtering it.

RELATED WORKS

The quality of CGI has become a very popular topic amongst academics and researchers (Antoniou & Skopeliti, 2015). Several critical literature reviews (or surveys) involving categorization have been conducted to provide an overview of this area (e.g. Bordogna et al., 2016; Mirbabaie et al., 2016; Senaratne et al., 2017; Wiggins et al., 2011).

An analysis of the quality assessment methods was carried out by Wiggins et al. (2011), where the authors analyzed the data validation policy and quality assessment in Citizen Science projects (i.e., Participatory Sensing). They found that the most common type of data validation is ex post, which is based on expert reviews conducted by trusted individuals or moderators. Bordogna et al. (2016) have also analyzed CGI in Citizen Science projects. They initially reviewed and categorized CGI projects, by analyzing the way they deal with CGI quality. This work also provided a classification scheme and a critical description of the strategies currently adopted to improve CGI quality. Bordogna et al. (2016) and Wiggins et al. (2011) conducted an important overview of quality assessment methods and made significant recommendations for improving CGI quality during a research project. However, these authors only discuss methods for quality assessment of CGI in Citizen Science projects and fail to take account of other CGI sources such as Collaborative Mapping and Social Media.

Senaratne et al. (2017) conducted a critical literature review of the existing methods to assess the quality of the main types of CGI: text, image, and map. This review examines methods that are based on theories and discussions in the literature, and provides examples of the practical applicability of all the different approaches. However, this is a traditional literature review and, as many researchers have pointed out, traditional reviews are prone to bias, i.e., authors may decide only to include studies with which they are familiar or which support their particular standpoint. In an attempt to avoid this kind of bias, we conducted a systematic literature review (SLR) to discover the methods employed to assess the quality of CGI. An SLR adopts a replicable, scientific, and transparent approach to locate the most significant literature about a given topic or discipline.

Mirbabaie et al. (2016) conducted a systematic literature review on CGI in disaster management. The main goal of this review was to provide information about the quality elements that are used, as well as the methods that are employed to measure these elements. They found that attributes such as 'accuracy' and 'consistency' are mainly used as criteria for quality assessment, while other factors such as 'trustworthiness' are not fully taken into account. However, they failed to conduct an in-depth analysis of the existing methods with regard to their applications and limitations and were only concerned with the existing methods for disaster management. To fill this gap, we carried out an SLR to discover the existing quality assessment methods for CGI in different application domains and discussed the limitations of each method.

METHODOLOGY

Systematic Literature Review (SLR) was first applied to support evidence-based medicine. SLR is a kind of secondary study that aims at identifying, analyzing and interpreting all the available evidence related to a research topic (Kitchenham & Charters, 2007). Differently from the usual process of literature review, a SLR is undertaken in a formal, rigorous and systematic way (Biolchini et al. 2005; Okoli & Schabram, 2010), i.e. in a way that is unbiased and (in a certain degree) repeatable (Kitchenham & Charters, 2007). With this methodology is possible to summarize existing evidence and identify any gaps in a research topic, and provide a framework for position new research activities (Kitchenham & Charters, 2007). Recently, SLR has been applied to the field of Geographic Information Science (GIS) for analyzing the current state of research, for instance, on the use of CGI for disaster management (Horita et al. 2013); concerning methodologies application and use cases of Twitter as a Location-based Social Network (Steiger et al. 2015); and on the use of CGI within natural hazard analysis (Klonner et al. 2016).

In this work, an SLR was carried out to discover current research on CGI within the scope of quality assessment. More specifically, each study was analyzed with regard to the method designed for assessing the quality of CGI. In conducting the SLR, we complied with the guidelines recommended by Kitchenham & Charters (2007). The SLR follows a sequence of well-defined steps, which comprises (i) planning the review, (ii) conducting the review and (iii) reporting the review.

Review Planning

An important activity of the planning phase is to draft a clear and concise Research Question (RQ) (Brereton et al., 2007; Okoli & Schabram, 2010), since it will be used as a guide for the entire SLR process. As the main goal of this work is to discover studies proposing methods for the quality assessment of CGI, the following RQ has been raised:

RQ) What are the methods used to assess the quality of CGI?

In a SLR, existing evidence to answer a RQ can be obtained by carrying out a search string on electronic databases. To build the search string, we first selected the main terms of our RQ, i.e. crowdsourced geographic information and quality assessment. We also identified synonyms for each term in order to maximize the number of returned studies. The synonyms for the former term were extracted from See et al. (2016). We applied the Boolean operator OR to join the synonyms of each term and the Boolean operator AND to join the main terms. The search string is shown in Table 1.

The identification of studies began with the search string being applied to 5 (five) electronic databases (Table 2). We particularly selected this set of electronic databases in order to maximize the number of studies, since a single database cannot find all the existing evidence concerning a research topic (Brereton et al., 2007). Moreover, we selected this set because of their relevance to the research field, i.e. these electronic databases index the main journals and conferences of the area. However, owing to the idiosyncrasy of each electronic database, we had to adjust the search string to them since there were few relevant studies in this field. Hence, we had to remove some synonyms of the search string (e.g., social computing, quality analysis and quality enhancement).

The inclusion and exclusion criteria assisted the selection of key studies that could be used to answer the research question (Biolchini et al., 2005; Petersen et al., 2007). Considering the main goal of this SLR and the aforementioned RQ, we defined a set of inclusion and exclusion criteria (Table 3) that were used as a basis for the selection of studies. Besides this set, a study was also excluded if (i) it is not published between 2004 and November 2015, (ii) it is not written in Portuguese or English, (iii) it is a SLR, (iv) it is not available and (v) it is duplicated or incomplete.

Table 1. Search string

(VGI OR "volunteered geographic information" OR "volunteered geographic data" OR "crowdsourced geographic information" OR "crisis mapping" OR "collaborative mapping" OR OpenStreetMap OR ((crowdsourcing OR "crowd sourcing" OR "crowd-sourcing" OR "crowd-sourced data" OR "user-generated content" OR "social media" OR Twitter OR Flickr OR "collective intelligence" OR "collective knowledge" OR "citizen based") AND (geographic OR spatial OR geotagged OR georeferenced))) AND ("quality assessment" OR "quality assurance" OR "quality evaluation" OR "quality control" OR reliability OR credibility OR trust OR accuracy)

Table 2. Electronic databases

Electronic Database	URL
IEEE Xplore	www.ieeexplore.ieee.org
ACM Digital Library	www.portal.acm.org
Web of Science	www.webofknowledge.com
Science Direct	www.sciencedirect.com
SCOPUS	www.scopus.com

Table 3. The inclusion and exclusion criteria employed for the qualitative review

IC1: the study sets out or adopts an approach for the quality assessment of crowdsourced geographic information
EC1: the study is related to quality assessment, but not to crowdsourced geographic information
EC2: the study is related to crowdsourced geographic information, but not to quality assessment
EC3: the study is not related to crowdsourced geographic information or quality assessment

Review Results

The search in the electronic databases resulted in a total of 555 primary studies included, after the duplicate studies had been removed (Figure 1). After reading the title and abstract, we identified 501 studies that are not directly related to the quality assessment of CGI. However, if the objective of the work was not clear in the abstract, the study was included for a more in-depth analysis. After this stage, we included 54 studies for a complete reading.

While the complete reading, the included studies were analyzed to determine if the study was indeed a candidate to answer the RQ. If the study satisfies at least one exclusion criterion, this was excluded. However, if doubts emerged with regard to include the study; the opinion of another reviewer was taken into account to decide about it. After a close analysis, 18 studies were found that discuss methods for assessing CGI quality (Table 4). In the following section, we describe each method in detail.

QUALITY ASSESSMENT METHODS

In this section, we describe the existing methods for quality assessment of CGI in the absence of authoritative data.

Geographic Context

Description: Each place has its own distinguishing characteristics. The basic idea entails investigating the area surrounding the CGI location to determine its characteristics and, make an assessment of its quality. Both physical and human characteristics should be taken into account when undertaking this.

Example: Senaratne et al. (2013) examined the geographic context around Flickr photographs to determine which areas can be viewed from the CGI location, or more specifically, whether the object can be viewed from the CGI location. Similarly, Zielstra & Hochmair (2013) carried out an investigation to evaluate the positional accuracy of geotagged photos from Flickr and Panoramio.

Redundancy of volunteers' contribution

Description: This method involves requesting several volunteers to provide information about the same geographic feature. Later, the CGI quality is determined by analyzing the volunteers' contributions by finding out whether or not there is a convergence of the information produced independently by different volunteers. For example, a geographic feature is labeled by several people to indicate a part of the land cover. The resulting set of labels is then analyzed and there is an estimate of the probability that the geographic feature belongs to a land cover category.

Example: Comber et al. (2013) used redundancy to estimate the reliability of non-expert volunteers. Similarly, See et al. (2013) employed this method to evaluate the accuracy and consistency of volunteers when labeling land cover and determining the human impact on the environment. Foody et al. (2013) applied redundancy to measure the accuracy of each volunteer when labeling images of land cover from satellite sensors. In contrast, Foody (2013) explored the redundancy of contributions to situations in which a large proportion of data is provided by poor sources and/or is incomplete.

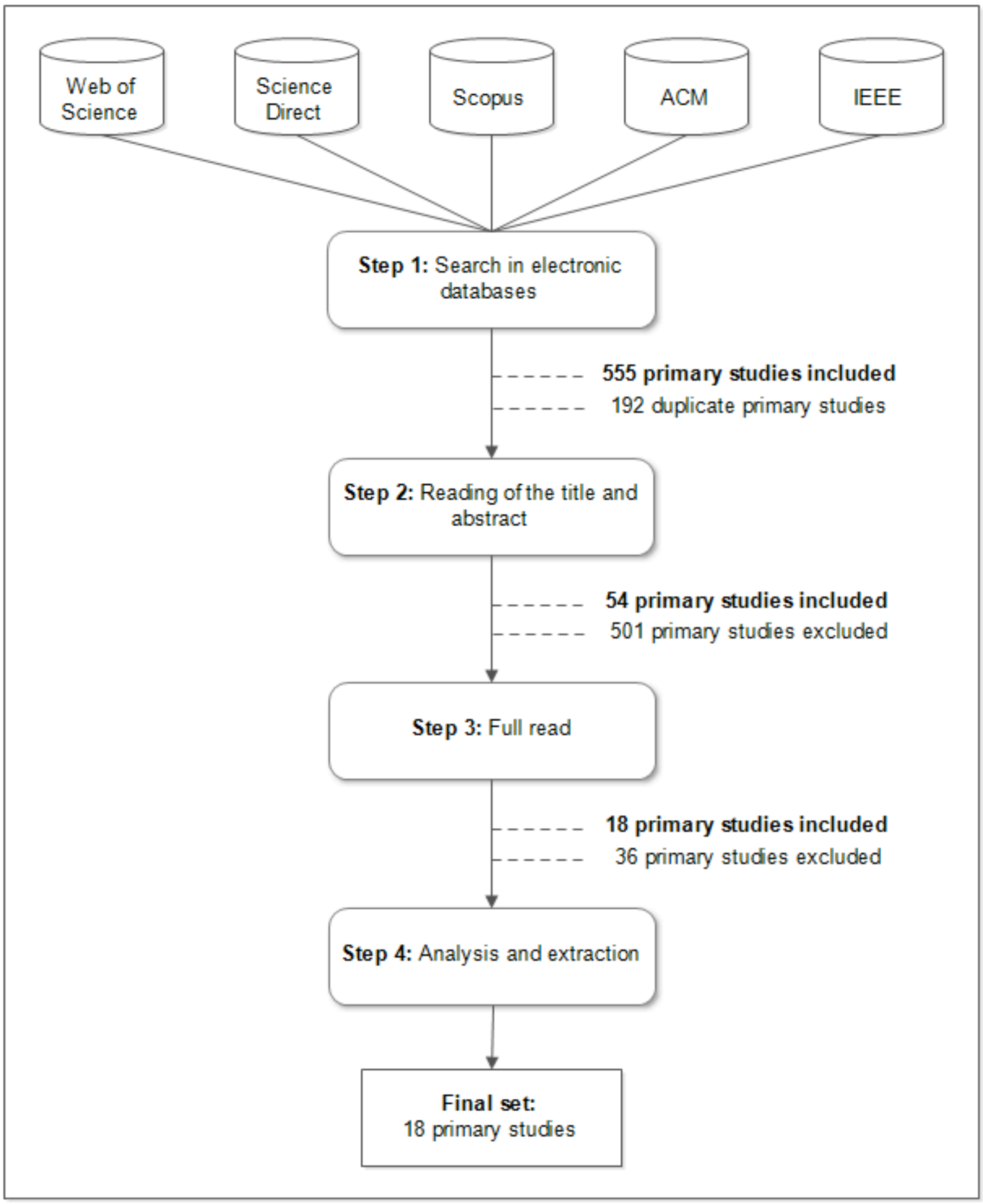

Figure 1. SLR process and the number of included and excluded studies in each step

Table 4. Detailed overview of the selected studies in the SLR

Study	Collaborative activity	Method	Quality element(s)	Objective
Karam & Melchiori (2013)	General	Ranking volunteers' contribution	Accuracy and completeness	Ranking volunteers' contribution to improve the accuracy and completeness of geo-spatial linked open data
Bordogna, Carrara, Criscuolo, Pepe, & Rampini (2014)	Participatory sensing	Ranking/filtering by linguistic terms	Fitness for purpose	Ranking and filtering high-quality CGI items in a glaciological citizen science project
Mohammadi & Malek (2015)	Collaborative mapping	Pattern extraction	Positional accuracy	Extracting patterns from corresponding reference data to estimate the positional accuracy of no corresponding reference OpenStreetMap data
Foody et al. (2013)	Participatory sensing	Redundancy of volunteers' contribution	Thematic accuracy	Employing redundancy of volunteers' contribution to evaluate the thematic accuracy of CGI
Cui (2013)	Participatory sensing	Automatic location checking	Spatial accuracy	Automatic checking of the positional accuracy of farmer markets' location
Jilani et al. (2014)	Collaborative mapping	Extraction/learning of characteristics	Semantic accuracy	Extraction and learning of geometrical and topological properties to assess the semantic accuracy of street network data from OpenStreetMap
Longueville et al. (2010)	Social media	Spatiotemporal clustering	Credibility	Spatiotemporal clustering of CGI items on social media to assess its credibility
See et al. (2013)	Participatory sensing	Redundancy of volunteers' contribution	Accuracy	Employing redundancy of volunteers' contribution to assess the accuracy of crowdsourced data on land cover
Ali & Schmid (2014)	Collaborative mapping	Data-based inference	Plausibility	Data-based inference to predict the correct class of a new entity in OpenStreetMap

Table 4. Detailed overview of the selected studies in the SLR

Study	Collaborative activity	Method	Quality element(s)	Objective
Haklay, Basiouka, Antoniou, & Ather (2010)	Collaborative mapping	Error detection/correct by crowd	Positional accuracy	Error detection and correction by crowd to improve positional accuracy of road network from OpenStreetMap
Bodnar et al. (2014)	Social media	Volunteer's profile; reputation	Trustworthiness	Analyzing volunteer's profile to determine the trustworthiness of his/her posts on social media
Zielstra & Hochmair (2013)	Social media	Geographic context	Positional accuracy	Analysis of the geographic context to evaluate the positional accuracy of Flickr and Panoramio photos
Foody (2013)	Participatory sensing	Redundancy of volunteers' contribution	Thematic accuracy	Exploring the redundancy of volunteer's contribution to assess the thematic accuracy of CGI in situations where a large proportion of data might come from spammers and/or be incomplete
Lertnattee et al. (2010)	Participatory sensing	Scoring volunteers' contribution	Reliability	Scoring CGI items in order to assess their reliability
Bishr & Kuhn (2013)	Participatory sensing	Volunteer's profile; reputation	Trustworthiness	Analysis of volunteer's reputation to evaluate CGI quality regarding the quality of water from a well
Keßler & de Groot (2013)	Collaborative mapping	Historical data analysis	Trustworthiness	Analysis of historical data to assess the trustworthiness of OpenStreetMap data in Muenster, Germany
Comber et al. (2013)	Participatory sensing	Redundancy of volunteers' contribution	Reliability	Employing redundancy of volunteers' contribution to determine the reliability of VGI with regard to the type of land cover
Senaratne et al. (2013)	Social Media	Geographic context	Positional accuracy	Analysis of the geographic context to assess the positional accuracy of Flickr photos

Scoring volunteers' contribution

Description: In crowdsourcing-based platforms, volunteers share their knowledge and opinions with the public. In this method, every contribution made by volunteers is awarded a score which is attached to the information. This score represents the "content quality" and it can be measured by means of different techniques.

Example: Lertnattee et al. (2010) measured the score by counting the number of votes, i.e. giving a voting score. Thus, information in a higher position in the hierarchy tends to be more believable by the users.

Ranking volunteers' contribution

Description: Feedback and contributions from other volunteers, that are familiar with an area, greatly assist in estimating the quality of CGI (Karam & Melchiori, 2013). In this method, CGI is submitted to experts who have more knowledge about the geographic area and are responsible for checking the quality of CGI and change, i.e. to correct it, if necessary. Later, all the CGI are ranked on the basis of the changes made by the experts and the historical record of activities carried out by a volunteer.

Example: Karam & Melchiori (2013) ranked CGI by employing four metrics: (i) the historical records of activities carried out by a volunteer, (ii) the number of activities carried out by other volunteers, (iii) the feedback received after the change was made and (iv) the scoring of the user that submitted the information.

Automatic location checking

Description: The location of a volunteer can be ascertained in different ways, such as GPS (Global Positioning System), manual georeferencing, or an address. However, the last of these could contain errors or typos. A way of automatically estimating it is to compare geocoded coordinates, from multiple geocoding services, with each other. Hence, the address should be submitted to several geocoding services, which will result in coordinates that are represented by latitude and longitude, and then provide the resulting address. To qualify as reference data, the different geocoding services should yield concordant results within a certain distance (Cui, 2013). The quality is, thus, measured as the distance between the geocoded coordinate and the actual coordinate.

Example: Cui (2013) employed automatic checking to determine the spatial accuracy of the location of a farmers' market.

Spatiotemporal clustering

Description: CGI quality can be addressed by aggregating information from several volunteers (Mummidi & Krumm, 2008) and, later, by evaluating the significance of the resulting clusters for a specific purpose (Longueville et al., 2010). Thus, instead of checking the quality of a single CGI element, the elements are evaluated as a whole, i.e. the quality of the CGI cluster is assessed.

This method consists of creating spatiotemporal clusters of CGI elements using prior information about a phenomenon of interest. The clusters are created on the basis of the assumption that "CGI elements created at the same place and time refer to the same event" (Longueville et al., 2010).

The process starts by creating temporal clusters, which are the CGI elements clustered in several temporal classes. After this, the temporal classes are divided into sub-classes in compliance with spatial criteria. These steps convert raw CGI of an unknown quality into spatiotemporal clusters, the importance of which can be quantified by means of a ranking score, which reflects the likelihood that an event took place in the time period and area that each cluster refers to.

Example: Longueville et al. (2010) clustered CGI to derive the likelihood that a flood event took place..

Volunteer's profile; reputation

Description: The volunteer is an important factor in the quality of CGI since his/her knowledge and background can have an influence on it. According to Fava (2015), for instance, expert volunteers provide higher-quality information than non-expert volunteers. Thus, they can act as a marker for the quality of their contributions. By employing this method, the volunteer's profile or reputation can be analyzed and used as a basis to estimate the quality of CGI. Bishr & Janowicz (2010) argue that if the volunteer has a reputation for being trustworthy; this means that his or her contribution is trustworthy too.

Example: Bodnar et al. (2014) employed this method to establish the veracity of four events, two of which turned out to be hoaxes, that took place in the United States. In contrast, Bishr & Kuhn (2013) employed this method to assess the trustworthiness of volunteers' statements regarding the quality of water from a well.

Error detection/correct by crowd

Description: According to Linus's Law, "given enough eyeballs, all bugs are shallow" (Raymond, 1999 *apud* Haklay, Basiouka, Antoniou, & Ather, 2010). In the case of CGI, this can be understood as 'given enough volunteers, (almost) all errors can be identified and corrected'.

The basic idea behind this method is that a single individual might unintentionally introduce an error in a crowdsourcing-based platform. Later, other people might notice this error and correct it, and hence the community of volunteers acts as gatekeepers.

Example: Haklay, Basiouka, Antoniou, & Ather (2010) investigated whether Linus's Law applies to the positional accuracy of OSM data.

Data-based inference

Description: Each geographic feature has its own characteristics (i.e. shape, size, etc.) which can be used to classify it. This method consists of learning the characteristics of each geographic feature and, later, using them to infer the correct type of a new entity. The inference is carried out by noting similarities with existing entities. The method can also be used to detect an incorrect classification of geographic features.

Example: Ali & Schmid (2014) designed a classifier that learns the correct class of existing entities (i.e. parks and gardens) on the basis of their characteristics (e.g. size) and used it to predict the correct class of a new entity.

Pattern extraction

Description: Assessing CGI quality without ground-truth data is a challenge. An alternative is to extract patterns from CGI with corresponding reference data and use the extracted pattern to estimate the quality of CGI without the need for a corresponding reference dataset. To achieve this goal, it is first necessary to define the indicators that can have an influence on the quality of CGI. Later, these can be used, together with the definition, to extract a pattern of the relation. Finally, the value of the defined indicators from the CGI with no corresponding dataset can be used with the patterns to estimate its quality.

Example: Mohammadi & Malek (2015) estimated the positional accuracy of no corresponding reference (NCR) OSM data by extracting a pattern from the corresponding reference (CR) OSM data.

Extraction/learning of characteristics

Description: Data representation can be regarded as the first stage towards knowledge discovery and allows the characteristics that are representative of each type of data to be extracted. Once the characteristics of a dataset have been obtained, the next stage entails learning the information implicit in the characteristics of the data.

Example: Jilani et al. (2014) extracted geometrical and topological properties of OSM street network data to infer the "road class" from the new data.

Ranking/filtering by linguistic terms

Description: The underlying principle of this method is the need to express the criteria linguistically. The linguistic terms are used to specify the desired values of the quality indicators and together, these comprise a schema for quality evaluation. Each CGI item is first evaluated on the basis of each criterion that is expressed linguistically and, later, ranked in degrees of global satisfaction. Finally, CGI items are filtered by being subject to the constraints of the application domain.

Example: Bordogna et al. (2014) employed this method to assess the quality of a CGI item (i.e. a picture) in a glaciological citizen science project.

Historical data analysis

Description: In special cases, CGI comes with historical data. In OSM, for instance, a new version of an object is created whenever its geometry is changed. From the history of the data, it is possible to derive (intrinsic) indicators that allow approximate statements to be made regarding data quality (Barron et al., 2014). An example of an indicator is the number of contributors since it has been demonstrated that this has an influence on the quality (Haklay et al., 2010). Moreover, the historical data can be analyzed to identify patterns and make predictions.

Example: Keßler & de Groot (2013) produced a set of indicators based on historical data (i.e., number of versions, contributors, confirmations, tag corrections and rollbacks) to assess the quality of OSM data in Muenster, Germany.

DISCUSSION AND CONCLUSION

CGI is obtained by means of collaborative mapping, participatory sensing and social media activities. Quality assessment is an important stage after gathering it since the information comes from unknown sources and is of unknown quality. This paper presents the results of an SLR that was carried out to discover the methods that can be employed to assess the quality of CGI in the absence of authoritative data. The search was carried out in five electronic databases to maximize the number of studies candidate to answer our RQ. After a close analysis, we found 13 methods to assess CGI quality when authoritative data is not available.

Unlike in existing studies (Section 3), we describe how each method works and what its limitations are. The limitations should be underlined because they could prevent the applicability of the method. We also classified the methods with regard to the type of collaborative activity. By doing this, we aim to highlight where each method has already been employed and, in a certain way, draw attention to areas where the method has not been applied yet.

While most of the methods have been used to assess the quality of CGI from Participatory Sensing, fewer methods were designed to assess the quality of CGI from Social Media. This can be explained by the fact that assessing the quality of this type of CGI is still a challenge due to the variety of ways the individuals communicate. However, Social Media has become a valuable source of information in the past few years since people use it share opinions and knowledge with a contact network. Thus, future studies should focus on other potential methods to assess the quality of CGI from Social Media.

Our analysis also revealed that some methods are not entirely independent of authoritative data, i.e. part of CGI dataset must have the corresponding authoritative data. One possible explanation is the lack of evidence that proves the efficiency of intrinsic analysis. Other methods rely on CGI's metadata. However, these methods could be affected because of missing metadata or privacy issues.

Although SLR is a rigorous and systematic methodology, there are some threats to its validity. These have been avoided by selecting several synonyms for both the main keywords with the aim of discovering all the primary studies in the area. However, we had to exclude some of them because we were hampered in our attempt to find relevant studies by faults/failures in the different search engines employed by each electronic database. In addition, the number of studies included might have been affected by language restrictions, as only studies written in English and Portuguese were taken into account. Thus, it is possible that some relevant studies were not included in this work.

ACKNOWLEDGMENTS

The authors would like to express thanks for the financial support provided by CAPES. Lívia Castro Degrossi is grateful for the financial support from CNPq (Grant no. 201626/2015-2) and CAPES (Grant No. 88887.091744/2014-01). João Porto de Albuquerque acknowledges financial support from CAPES (Grant no. 88887.091744/2014-01), and Heidelberg University (Excellence Initiative II/Action 7). Roberto dos Santos Rocha is grateful for the financial support from European Community's Seventh Framework Programme: Marie Curie Actions/Initial Training Networks (Grant no. 317382) and Federal University of Bahia.

REFERENCES

Albuquerque, J. P. de, Herfort, B., Eckle, M., & Zipf, A. (2016). Crowdsourcing geographic information for disaster management and improving urban resilience: an overview of recent developments and lessons learned. In C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, & R. Purves (Eds.), *European Handbook of Crowdsourced Geographic Information* (pp. 309–321). Ubiquity Press.

Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2014). A Geographic Approach for Combining Social Media and Authoritative Data towards Identifying Useful Information for Disaster Management. *International Journal of Geographical Information Science*. http://doi.org/http://dx.doi.org/10.1080/13658816.2014.996567

Ali, A. L., & Schmid, F. (2014). Data Quality Assurance for Volunteered Geographic Information. In *Geographic Information Science* (Vol. 8728, pp. 126–141). Springer International Publishing. http://doi.org/10.1007/978-3-319-11593-1_9

Antoniou, V., & Skopeliti, A. (2015). Measures and Indicators of Vgi Quality: an Overview. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-3/W5*(September), 345–351. http://doi.org/10.5194/isprsannals-II-3-W5-345-2015

Ballatore, A., & Zipf, A. (2015). A Conceptual Quality Framework for Volunteered Geographic Information. In *Spatial Information Theory: 12th International Conference, COSIT 2015, Santa Fe, NM, USA, October 12-16, 2015, Proceedings* (pp. 89--107). Springer International Publishing. http://doi.org/10.1007/978-3-319-23374-1_5

Barron, C., Neis, P., & Zipf, A. (2014). A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS*, *18*(6), 877–895. http://doi.org/10.1111/tgis.12073

Biolchini, J., Mian, P. G., Candida, A., & Natali, C. (2005). *Systematic Review in Software Engineering*. *Engineering*. Rio de Janeiro, Brazil.

Bishr, M., & Janowicz, K. (2010). Can we Trust Information ? - The Case of Volunteered Geographic Information. In *Proceedings of the Workshop Towards Digital Earth: Search, Discover and Share Geospatial Data*. Berlin, Germany.

Bishr, M., & Kuhn, W. (2013). Trust and Reputation Models for Quality Assessment of Human Sensor Observations. In *Spatial Information Theory* (pp. 53–73). Springer International Publishing. http://doi.org/10.1007/978-3-319-01790-7_4

Bodnar, T., Tucker, C., Hopkinson, K., & Bilen, S. G. (2014). Increasing the veracity of event detection on social media networks through user trust modeling. In *IEEE International Conference on Big Data (Big Data)* (pp. 636–643). IEEE. http://doi.org/10.1109/BigData.2014.7004286

Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M., & Rampini, A. (2014). A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science. *Information Sciences*, *258*, 312–327. http://doi.org/10.1016/j.ins.2013.07.013

Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M., & Rampini, A. (2016). On predicting and improving the quality of Volunteer Geographic Information projects. *International Journal of Digital Earth*, *9*(2), 134–155. http://doi.org/10.1080/17538947.2014.976774

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. http://doi.org/10.1016/j.jss.2006.07.009

Comber, A., See, L., Fritz, S., Velde, M. Van Der, Perger, C., & Foody, G. (2013). Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied Earth Observation and Geoinformation*, *23*(1), 37–48. http://doi.org/10.1016/j.jag.2012.11.002

Cui, Y. (2013). A systematic approach to evaluate and validate the spatial accuracy offarmers market locations using multi-geocoding services. *Applied Geography*, *41*, 87–95. http://doi.org/10.1016/j.apgeog.2013.03.016

Degrossi, L. C., Albuquerque, J. P. de, Fava, M. C., & Mendiondo, E. M. (2014). Flood Citizen Observatory: a crowdsourcing-based approach for flood risk management in Brazil. In *Proceedings of the 26th International Conference on Software Engineering and Knowledge Engineering*.

Fava, M. C. (2015). *Modelo de Alerta Hidrológico com Base Participativa usando Sistema de Informações Voluntárias para Previsão de Enchentes*. University of São Paulo.

Foody, G. M. (2013). Rating crowdsourced annotations: evaluating contributions of variable quality and completeness. *International Journal of Digital Earth*, *0*(0), 1–21. http://doi.org/10.1080/17538947.2013.839008

Foody, G. M., See, L., Fritz, S., der Velde, M. Van, Perger, C., Schill, C., & Boyd, D. S. (2013). Assessing the Accuracy of Volunteered Geographic Information arising from Multiple Contributors to an Internet Based Collaborative Project. *Transactions in GIS*, *17*(6), 847–860. http://doi.org/10.1111/tgis.12033

Foody, G. M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., & Boyd, D. S. (2013). Assessing the

Accuracy of Volunteered Geographic Information arising from Multiple Contributors to an Internet Based Collaborative Project. *Transactions in GIS*, *17*(6), 847–860. http://doi.org/10.1111/tgis.12033

Girres, J. F., & Touya, G. (2010). Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, *14*(4), 435–459. http://doi.org/10.1111/j.1467-9671.2010.01203.x

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211–221. http://doi.org/10.1007/s10708-007-9111-y

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, *3*(3), 231–241. http://doi.org/10.1080/17538941003759255

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, *1*, 110–120.

Haklay, M., Basiouka, S., Antoniou, V., & Ather, A. (2010). 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), 315–322. http://doi.org/10.1179/000870410X12911304958827

Harvey, F. (2013). To Volunteer or to Contribute Locational Information? Towards Truth in Labeling for Crowdsourced Geographic Information. In *Crowdsourcing Geographic Knowledge* (pp. 31–42). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-007-4587-2_3

Horita, F. E. A., Degrossi, L. C., Assis, L. F. F. G., Zipf, A., & Albuquerque, J. P. de. (2013). The use of Volunteered Geographic Information and Crowdsourcing in Disaster Management: a Systematic Literature Review. In *Proceedings of the Nineteenth Americas Conference on Information Systems*.

Jilani, M., Corcoran, P., & Bertolotto, M. (2014). Automated highway tag assessment of OpenStreetMap road networks. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 449–452).

Karam, R., & Melchiori, M. (2013). A crowdsourcing-based framework for improving geo-spatial open data. In *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 468–473). http://doi.org/10.1109/SMC.2013.85

Keßler, C., & de Groot, R. T. A. (2013). Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap. In *Geographic Information Science at the Heart of Europe* (pp. 21–37). Springer International Publishing. http://doi.org/10.1007/978-3-319-00615-4_2

Kitchenham, B. A., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*. Keele, UK.

Klonner, C., Marx, S., Usón, T., Porto de Albuquerque, J., & Höfle, B. (2016). Volunteered Geographic Information in Natural Hazard Analysis: A Systematic Literature Review of Current Approaches with a Focus on Preparedness and Mitigation. *ISPRS International Journal of Geo-Information*, *5*(7), 103. http://doi.org/10.3390/ijgi5070103

Lertnattee, V., Chomya, S., & Sornlertlamvanich, V. (2010a). Reliable Improvement for Collective Intelligence on Thai Herbal Information. In *Advances in Intelligent Information and Database Systems* (pp. 99–110). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-12090-9_9

Lertnattee, V., Chomya, S., & Sornlertlamvanich, V. (2010b). Reliable Improvement for Collective Intelligence on Thai Herbal Information. *Studies in Computational Intelligence*, *283*, 99–110.

Longueville, B. De, Luraschi, G., Smits, P., Peedell, S., & Groeve, T. De. (2010). Citizens as Sensors for Natural Hazards: A VGI Intergration Workflow. *Geomatica*, *64*, 41–59.

Mirbabaie, M., Stieglitz, S., & Volkeri, S. (2016). Volunteered Geographic Information and its Implications for Disaster Management. In *Proceedings of the 49th Hawaii International Conference on System Sciences*.

Mohammadi, N., & Malek, M. (2015). Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data. *Journal of Spatial Science*, *60*(1), 119–135. http://doi.org/10.1080/14498596.2014.927337

Mooney, P., Corcoran, P., & Winstanley, A. C. (2010). Towards quality metrics for OpenStreetMap. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems GIS 10*, 514–517. http://doi.org/10.1145/1869790.1869875

Mummidi, L., & Krumm, J. (2008). Discovering points of interest from users' map annotations. *GeoJournal*, *72*(3), 215–227.

Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information

Systems Research. *All Sprouts Content*. Retrieved from http://aisel.aisnet.org/sprouts_all/348

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2007). Systematic Mapping Studies in Software Engineering. In *12th international conference on Evaluation and Assessment in Software Engineering* (pp. 68–77). Italy: British Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=2227115.2227123

Raymond, E. S. (1999). *The Cathedral and the Bazaar*. O'Reilly Media.

See, L., Comber, A., Salk, C., Fritz, S., van der Velde, M., Perger, C., … Obersteiner, M. (2013). Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts. *PLoS ONE*, *8*(7), 1–11. http://doi.org/10.1371/journal.pone.0069958

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., … Rutzinger, M. (2016). Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. *ISPRS International Journal of Geo-Information*, *5*(5), 55. http://doi.org/10.3390/ijgi5050055

Senaratne, H., Bröring, A., & Schreck, T. (2013). Using reverse viewshed analysis to assess the location correctness of visually generated VGI. *Transactions in GIS*, *17*(3), 369–386. http://doi.org/10.1111/tgis.12039

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, *31*(1), 139–167. http://doi.org/10.1080/13658816.2016.1189556

Standardization, I. O. for. (2013). *Geographic information - Data quality (ISO 19157)*.

Steiger, E., de Albuquerque, J. P., & Zipf, A. (2015). An Advanced Systematic Literature Review on Spatiotemporal Analyses of Twitter Data. *Transactions in GIS*, *19*(6), 809–834. http://doi.org/10.1111/tgis.12132

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234. http://doi.org/10.2307/143141

Wiggins, A., Newman, G., Stevenson, R. D., & Crowston, K. (2011). Mechanisms for Data Quality and Validation in Citizen Science | Kevin Crowston. In *IEEE eScience Conference Computing for Citizen Science*. Stockholm, Sweden. Retrieved from https://crowston.syr.edu/node/422

Zielstra, D., & Hochmair, H. H. (2013). Positional accuracy analysis of Flickr and Panoramio images for selected world regions. *Journal of Spatial Science*, *58*(2), 251–273. http://doi.org/10.1080/14498596.2013.801331