

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/322705917>

# Volunteered geographic information quality assessment using trust and reputation modelling in land administration systems in developing countries

Article in *International Journal of Geographical Information Science* · January 2018

DOI: 10.1080/13658816.2017.1409353

CITATIONS

25

READS

375

3 authors:



**Kealeboga Kaizer Moreri**

University of Botswana

21 PUBLICATIONS 101 CITATIONS

[SEE PROFILE](#)



**David Fairbairn**

Newcastle University

90 PUBLICATIONS 1,130 CITATIONS

[SEE PROFILE](#)



**Philip Michael James**

Newcastle University

96 PUBLICATIONS 1,948 CITATIONS

[SEE PROFILE](#)

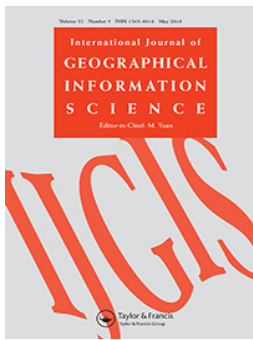
Some of the authors of this publication are also working on these related projects:



Botswana COVID-19 Spatial Mapping and Web Publication [View project](#)



VGI data [View project](#)



## Volunteered geographic information quality assessment using trust and reputation modelling in land administration systems in developing countries

Kealeboga K. Moreri, David Fairbairn & Philip James

To cite this article: Kealeboga K. Moreri, David Fairbairn & Philip James (2018) Volunteered geographic information quality assessment using trust and reputation modelling in land administration systems in developing countries, International Journal of Geographical Information Science, 32:5, 931-959, DOI: [10.1080/13658816.2017.1409353](https://doi.org/10.1080/13658816.2017.1409353)

To link to this article: <https://doi.org/10.1080/13658816.2017.1409353>



Published online: 25 Jan 2018.



Submit your article to this journal [↗](#)



Article views: 138



View related articles [↗](#)



View Crossmark data [↗](#)



RESEARCH ARTICLE



# Volunteered geographic information quality assessment using trust and reputation modelling in land administration systems in developing countries

Kealeboga K. Moreri , David Fairbairn and Philip James

School of Civil Engineering and Geosciences, Newcastle University, Newcastle Upon Tyne, UK

## ABSTRACT

This article presents an innovative approach to establish the quality and credibility of Volunteered Geographic Information (VGI) such that it can be considered in Land Administration Systems (LAS) on a Fit for Purpose (FFP) basis. A participatory land information system can provide affordable and timely FFP information about land and its resources. However, the establishment of such a system involves more than just technical solutions and administrative procedures: many social, economic and political aspects must be considered. Innovative approaches like VGI can help address the lack of accurate, reliable and FFP land information for LAS, but integration of such sources relies on the quality and credibility of VGI. Verifying volunteer efforts can be difficult without reference to ground truth: a novel Trust and Reputation Modelling methodology is proposed as a suitable technique to effect such VGI data set validation. This method has been applied to successfully demonstrate that VGI can produce accurate and reliable data sets which can be used to conduct regular systematic updates of geographic information in official systems. It relies on a view that the public can police themselves in establishing proxy measures of VGI quality thus facilitating VGI to be used on a FFP basis in LAS.

## ARTICLE HISTORY

Received 31 July 2017

Accepted 21 November 2017

## KEYWORDS

Volunteered geographic information; trust and reputation modelling; thematic accuracy; semantic accuracy; volunteer reputation

## 1. Introduction

This research has been motivated by challenges for Land Administration Systems (LAS) in developing countries, in particular a lack of regular updates and maintenance of geographic information. This leads to inefficiencies in the administering of land in these countries. Limited maintenance budgets prevalent in developing countries make it difficult for organizations to conduct regular systematic updates of geographic information. Despite these challenges, geographic information still forms a major component of effective LAS. For a LAS to remain useful, it must reflect realities on the ground and this can only be achieved if land information is reported regularly (Zevenbergen 2002). Biraro *et al.* (2015) stress that if changes in land are not captured in properly administered land registers, LAS lose societal relevance and are eventually replaced by informal systems. Current official systems are based on frameworks that are closed, expensive and prone to abuse. Land information held in current systems

may not be readily available to local communities, and a lack of openness can make it easy for officials to manipulate records for their interest or those of the elite and politically well connected. It is proposed that an alternative geospatial data collection mechanism that is affordable, participatory, transparent and inclusive of all stakeholders is the answer to land administration challenges in developing countries, especially in Africa.

The increase in collaborative initiatives like Wikipedia, OpenStreetMap (OSM) and Wikimapia is a positive sign that communities around the world are eager to share content of all types online using contemporary technologies and systems. Furthermore, the increase in the use of the Web, Global Positioning Systems (GPS) enabled smart phones, and wider and affordable internet access in developing countries, has greatly facilitated collaborative efforts among citizens. These efforts have opened doors for the public to become collectors of geographic information known as Volunteered Geographic Information (VGI) (Goodchild, 2009). VGI is a type of geographic information where people (experienced or inexperienced) either as individuals or collectively, voluntarily collect, organize and disseminate geographic information in Web-based environments (Tulloch 2008).

VGI initiatives like OSM and Wikimapia have shown that detailed geographic information can be provided in a timely and low-cost manner (Goodchild 2009, Goodchild and Li 2012, Fonte *et al.* 2015). However, they can suffer from serious weaknesses, such as lack of metadata about contributed data sets to inform potential reusers of the parameters employed for quality assurance measures (Goodchild and Li 2012). VGI quality assessment, among other aspects is not conducted in most projects due to a lack of central coordination and strict data collection guidelines (Haklay *et al.* 2010, Corcoran and Mooney 2013). These pose specific challenges when VGI is considered for adoption and incorporation in official systems.

VGI can support basic tasks like map production and updating (Moreri *et al.* 2015), but the involvement of volunteers, who in most cases are untrained or non-experts in handling geographic information, implies that VGI can be of varying quality. In VGI initiatives, people can collect geographic information without any guidance or instructions, leading to inconsistencies in the data collected. These challenges are further increased by the lack of systematic and comprehensive VGI quality assurance measures integral to geospatial data collection (Haklay *et al.* 2008). Thus, VGI is characterized by unstructured, heterogeneous, unreliable data which makes data integration for value-added purposes difficult to effect. These VGI quality challenges make land authorities reluctant to incorporate the contributed data sets into their official databases. Therefore, a Trust and Reputation Modelling (TRM) methodology is proposed to establish a 'proxy' quality and credibility measure of VGI without the typical reference to ground truth which characterizes most quality assessments. TRM utilizes the 'power of the crowd' principle (Haklay *et al.* 2010) to establish the level of trust of VGI and characterize the credibility of volunteers.

The 'power of the crowd' principle has proved successful in non-spatial collaborative initiatives like Wikipedia and open source software design. The rationale behind the principle is that inaccuracies in contributions are likely to be identified and corrected by many participants, thus reducing the errors. As this principle has not been explored extensively in VGI, this study has examined the potential of intrinsic measures of VGI

quality based on TRM, and applied them to data handling in LAS, particularly where authoritative data sets for ground truthing are limited.

Consensus-based decision-making (Collins and Mitchell 2017) is widely practised in rural areas of Africa for purposes such as land adjudication. When communities have a collective say, there exists a common sense of ownership and responsibility for stewardship. However, consensus-based decisions do have challenges: they can suffer from a lack of accountability from community members, especially if it is a large group without common goals, clearly implemented processes and active facilitators. A lack of clear decision-making processes can promote mistrust among community members. Moreover, community development can suffer, and disputes occur, when a consensus cannot be reached. Desirable features for good governance as outlined by FAO (2007) include as follows: (a) enabling citizens to participate fully in governance through consensus-building and engaging them without curbing their freedom of expression; (b) designing responsive systems that citizens want and need and (c) delivering quality services in the most effective and efficient way. Unfortunately, these features are currently lacking in LAS in developing countries.

Within the Open Source community, the assumption that, as the number of contributors increases so does the quality, is known as 'Linus' Law' (Raymond 2001). Haklay *et al.* (2010), have proved that this rule applies when assessing the positional accuracy of spatial features in VGI. This article proposes the application of this rule in determining the attribute and positional accuracy of VGI using TRM for land administration. As multiple participants work in the same geographic area, often capturing the same data, there are opportunities for errors made by others to be identified and fixed, consequently, improving the quality of the contributed data sets without the need for formal quality assurance measures. Practical studies of multiple annotations obtained from crowdsourcing activities reveal that high accuracies can be achieved from a small number of contributors (Snow *et al.* 2008, Haklay *et al.* 2010, Foody *et al.* 2015). The term 'crowdsourcing' is commonly applied in describing projects which use large numbers of contributors: in VGI this does not necessarily imply multiple collection of the same data, although in this article, this 'many eyes' principle is taken as the definition of crowdsourcing.

After a review of non-spatial and spatial initiatives that use the TRM concept, this article explores their successes and shortcomings in assessing quality in the products they serve; the TRM methodology, how it can be applied, and its implications are then outlined; a further section discusses the application of the research methodology in a case study related to land administration; and the results obtained are discussed and analysed.

## 2. TRM in non-spatial and spatial initiatives

TRM has been used in the past as a quality matrix for websites and Web services (Adler and De Alfaro 2007, Bishr and Kuhn 2007, Javanmardi *et al.* 2010). Several studies have been conducted in the past to confirm the practical effect that reputation has on Web-based activity, especially in building trust between online communities, as in e-commerce (e.g. eBay, Amazon) and for open, online encyclopaedia (Wikipedia) websites. An investigation of the successes and challenges of these initiatives was conducted to

inform the establishment of quality and credibility measures of contributed data sets in the context of land administration.

## 2.1. TRM in non-spatial initiatives

Prior to their introduction in geospatial initiatives, trust and reputation models have been key factors in the successful adoption and utilization of e-commerce websites (Sabater and Sierra 2005). The models have developed 'proxy' quality and credibility measures of goods and sellers, respectively, for example, using reputation to reflect the trustworthiness of individuals in online marketplaces (Zacharia *et al.* 1999, Mui 2002). Online e-commerce websites, like eBay and Amazon, use reputation as a function of the cumulative positive, neutral and negative ratings for sellers and buyers in regards to their transaction history (Resnick *et al.* 2000, Bajari and Hortacsu 2003).

Bajari and Hortacsu (2003) confirmed the empirical effect that reputation has on the Web, especially in building trust between buyers and sellers. One ongoing problem with online reputation systems is that they do not guard against the creation of pseudonyms: they can be prone to abuse and malicious attacks by individuals using false names, who can be mischievous yet unaffected by reputational consequences. Further challenges include imbalance between positive and negative feedback when quantifying reputation, lack of context dependence, and impact of reciprocity (e.g. mutual exchange of favour or revenge). Despite such challenges, reputation systems have seen tremendous growth in establishing the credibility and reliability of online sellers and the products they sell (Sonja *et al.* 2009), especially when initially 'setting up'.

In social websites, a rating system is used to allow users to express their level of agreement or disagreement with other user's contributions. There are many scales of measuring the quality of features in such rating systems, including unary scales, binary scales, and common five-star rating scales. Unary scales are popular in social networking sites, like Facebook, involving users clicking on a 'Like' button to show appreciation of content contributed by other users. The unary scale provides a single response value which may be insufficient in the context of contributed VGI content, where several parameters might be included. For geospatial content, such parameters can include its currency, geometric (positional), semantic (attribute) accuracy and completeness. These parameters would influence the final subjective judgment and award of a rating value for the geospatial entity.

Binary scales are popular with social news and video websites like YouTube, where users can express their like or dislike of a video by clicking the 'Thumbs Up' and 'Thumbs Down' icons, respectively. Such a system, like the unary system, is also not ideal for VGI, because it has a limited depiction of agreement or disagreement with a contributed entity.

A five-star rating scale is commonly used in recommender systems for commercial retailer stores like Amazon and eBay. It is used to highlight the trustworthiness of online buyers and sellers in the e-commerce websites. However, Cowan (2013) argues that the approach only allows products to be judged on their popularity and not on their details. Nonetheless, Babbie (2007) argues that a five-star rating scale, when used as a data reduction tool, allows a summary of several indicators into a single numerical value that could be assigned to an entity (including geographic entity) as its final quality value.

To address the challenge of a quantitative system offering a one-dimensional rating, a qualitative component has been introduced by eBay in the form of a commentary feedback. This provides more information to a potential customer, who tries to distinguish between two sellers with similar quantitative scores. In addition, it helps justify the quantitative scores given to a seller to provide more weight to the scores awarded (Kwan and Ramachandran 2009).

Wikipedia uses a content-driven reputation system where authors are evaluated based on how their contribution fares in the website (Adler and De Alfaro 2007). For example, when author A contributes an article to Wikipedia and author B revises it, she may choose to preserve some of author A's contributions, thus providing a vote of confidence in the contributions. Therefore, the reputation of author A will be increased based on the number of preserved contributions, as well as on the reputation of author B herself (Adler and De Alfaro 2007). Furthermore, the reputation of an individual in Wikipedia increases if their contributions and edits are persistent. However, if their content is revised quickly, this could affect their reputation negatively. Such approaches compute assessments of data quality based on data provenance (data origin), which is eventually combined with user feedback (Artz and Gil (2007)). Wikipedia entries are associated with an Internet Protocol (IP) address and a user account. This removes the anonymity element and facilitates the tracking of all edits and contributions made to the platform. In addition, it helps users to search for trusted content, by using IP addresses, especially of known organizations, to make better informed decisions.

TRM has played a major role in the growth of non-spatial initiatives, building trust between buyers and sellers in e-commerce websites. Moreover, TRM has been used to encourage honest online transactions by penalizing dishonest behaviour through loss of reputation. Despite challenges of abuse, malicious attacks and dishonesty, these initiatives continue to evolve, by developing security mechanisms to curb against attacks. Such experiences could be applied to VGI initiatives which currently do not have established standard measures of assessing quality.

## 2.2. TRM in spatial initiatives

In the spatial community, the concept of TRM has been applied to collaborative initiatives incorporating VGI approaches. Haklay *et al.* (2010) testing of Linus' Law to analyse data provenance in OSM data adopted the 'many eyes' principle. The analysis was mainly on the rollbacks and history of edits of contributed data sets which were then used to determine the quality and currency of the data sets, thus informing their FFP and reuse capabilities. In the context of spatial data systems, the FFP concept is not necessarily standards based, like traditional mapping projects, but rather it is concerned with more pragmatic approaches by end users (Enemark *et al.* 2014). For commercial and official geographic information, associated metadata and quality metrics are commonly presented to assist in determining FFP: but for data sets where such enhancements are unavailable, including VGI, trust and reputation models have been proposed as 'proxies' for data quality (Bishr and Kuhn 2007, Kessler and De Groot 2013).

Traditional data quality measures are generally lacking in VGI environments (Haklay *et al.* 2010, Heipke 2010, Osterman and Spinsanti 2011) as they rely on sound reference data which may be unavailable in some developing countries. Quality is difficult to

determine with typically dynamic VGI data sets. This has motivated some in the geographic community to investigate alternative measures of assessing the quality of contributed data sets and the credibility of contributors in VGI environments. These community-based collaborative models use trust matrices to assess the quality of contributed data sets in VGI platforms. For example, they involve volunteers contributing geographic information and their peers given the opportunity to subjectively judge and rate the quality of the contributed contents. Components which can be used to measure VGI quality include (a) attributes and semantic rigour (including folksonomies); (b) positional accuracy and (c) volunteer reputation.

Related to each of these components, trust can be used as a measure of quality in a collaborative environment like VGI, adopted as a 'proxy' measure of geospatial information quality (Bishr and Kuhn 2007). A trust model developed by Bishr and Kuhn (2007) classified and filtered collaboratively contributed geographic information, relying on 'folksonomies' as a means of collecting metadata about user generated content a folksonomy is a collaboratively generated classification system, similar to an ontology, that enables users to categorize attributes they contribute or encounter on the Web (Golder and Huberman 2006). Trust here is measured subjectively, where a trust-rated entity is considered of 'satisfactory' quality if it is regarded as useful and relevant to a larger group of consumers.

Assessing quality by positional accuracy of VGI has attracted more interest from the geographic information research community (Haklay *et al.* 2010, Mooney *et al.* 2010, Fairbairn and Al-Bakri 2013). These studies have mainly compared contributed data with data sets in official databases. However, this approach suffers major drawbacks (D'antonio *et al.* (2014)), since it requires access to professional data sets, often expensive and/or unavailable. Moreover, the quality assessment procedures developed are not universally valid, especially in those areas where ground truth data sets are inaccessible.

To address these issues, D'antonio *et al.* (2014) proposed a model that evaluates a volunteer's reputation and data trustworthiness deriving information from VGI data itself, rather than comparing it with external sources. The model developed identifies basic editing types where feature versions are evaluated against three characteristics: semantic attributes, geometric properties and qualitative spatial relations (e.g. disjoint, overlap, contains, cover). Using these characteristics, they concluded that data trustworthiness and reputation are a function of their direct and indirect editing effects over a period of time. That is, as more edits are performed on a contributed entity over a period of time, then more errors can be identified and corrected to improve data quality.

VGI is authored by heterogeneous sources and therefore there is a need to establish mechanisms to assess the credibility of contributors. Certain characteristics of a contributor can be used as 'proxy' measures of reliability and inherent quality of the data sets they produce (Flanagin and Metzger 2008, Golbeck 2008). The reputation of a volunteer can be based on several personal aspects that involve their qualifications, experience in handling spatial data, activity space (Goodchild 2009) and their motivations to contribute to a VGI initiative (Flanagin and Metzger 2008). It is argued that a contributor in geographic proximity to a source can produce more current local information compared to those further away from it (Goodchild 2009).

Van Exel *et al.* (2010) suggest the use of three components for determining volunteer reputation: local knowledge, experience and recognition: (a) local knowledge helps



consumers identify missing or incorrect contributions relatively easily; (b) the experience of a volunteer in contributing to a project is correlated to their overall interaction with the system over time and the quality of their contributions and (c) recognition entails the awareness given to contributors by other consumers when a certain threshold is met. Volunteers may be biased on their contributions towards their preferred land parcels. Moreover, the subjective nature of VGI makes it difficult to filter and detect (or reject) bias on contributed data sets. Nonetheless, machine learning techniques may be employed to address this issue (Hagenauer and Helbich 2012, Ali and Schmid 2014).

Whether spatial or non-spatial, TRM systems are clearly embedded in the commercial arena, implicitly used in transactions and online data handling, can be applied to a range of data, and have potential in volunteer data handling projects: despite this, they do have shortcomings that need to be highlighted particularly in a VGI context.

### **2.3. Challenges of trust and reputation systems**

Reliable trust and reputation models have the potential to increase cooperation between contributors and consumers and thus improve the usability and overall performance of online applications (Mozhgan 2012). Such models are built around feedback and human interaction, but before they can be accepted as a legitimate trust solution, it is necessary to understand how they may be compromised and how subsequent problems can be addressed. Challenges that weaken TRM when assessing the quality of online entities and the credibility of participants can be grouped into four categories: (a) feedback generation, (b) feedback distribution, (c) feedback aggregation and (d) subjective feedback (Josang and Golbeck 2009, Mozhgan 2012).

*Feedback generation* involves users/consumers providing feedback to describe or rate their experiences in dealing with a system or entity. Online systems have developed rating systems to collect participant feedback, but these can be abused or ignored due to (a) the inability of the system to provide incentives to motivate participants to provide feedback; (b) bias by participants to provide positive feedback because of friendly actions and (c) cold start problems experienced by new volunteers. The latter occur when new volunteers find it difficult to raise their reputation score due to the reluctance of other individuals to deal with low reputation volunteers. eBay and Amazon are examples of organizations that use a feedback generation system to provide quality and credibility measures of products and sellers, respectively. Other challenges of feedback generation include the creation of pseudonyms, where participants create multiple profiles to initiate problematic behaviour like posting misleading information.

*Feedback distribution* involves collected feedback not being distributed comprehensively or appropriately to represent the entity being rated. For example, formerly, the reputation system of an eBay seller was based on a single measure which failed to distinguish whether the score provided by a buyer was awarded for the quality of products sold or the efficiency of the seller in delivering products on time.

*Feedback aggregation* occurs when the collected trust value of a participant is not representative of their past actions. For example, in eBay a seller with 20 successful sales and 5 unsuccessful ones will have an equal rating with a seller with only 15 successful sales and no unsuccessful ones. This is a challenge which occurs because of inaccurate algorithms, like the value imbalance equation which weighs all feedback equally

regardless of transaction value (Tavakolifard and Almeroth 2012). According to Dellarocas (2002), a participant can take advantage of this property to build a good reputation by executing small value trades and use the reputation accumulated to cheat in a high value transaction.

The fourth challenge is *subjective feedback* which is usually based on the personal taste and cultural background of a participant (Bishr and Mantelas 2008). What is viewed as good by one person may be viewed as bad by another.

The challenges highlight the subjective nature of feedback and how online networks can be compromised. They can limit the effectiveness of these networks preventing their use and consideration in official systems. However, despite the weaker security guarantees served by TRM systems, they have been applied successfully in many peer-to-peer online systems to establish trust between consumers and sellers. To expand the scope of traditional security models, trust-based systems have emerged as solutions for citizens to accept risks and deal with uncertainty.

## 2.4. Sample number considerations in establishing VGI quality

This study advances the ‘wisdom of the crowd’ principle (Raymond 2001), to establish the ‘proxy’ quality and credibility of VGI and its creators, those engaged in collaborative knowledge building processes. No agreement has yet been reached on the sufficient number of participants needed to establish VGI quality, to have a representative sample of the data items collected, to achieve acceptable results and to gauge participants’ reliability. Also, the number of samples can depend upon the entity image classification, vector alignment or positioning of point observations may each have different optimum sampling sizes. Successful collaborative mapping projects like OSM embody the collective intelligence philosophy to assemble user contributions into a ‘patchwork’ map (Spielman 2014). Through established review processes, OSM aggregates participants’ contributions into a single map for use by the spatial community at large. The trust and reputation concept as used in text-based initiatives, which relies on the number of contributors and volume of feedback to enforce quality is advanced here as the central tenet of a system to establish the overall quality of VGI. The basis of TRM is that the more participants engage in the initiative the better, as more bugs can easily be identified and fixed. Spielman (2014) confirmed that the more users and contributors a geospatial community has, the more likely it is to produce better quality products.

Mooney *et al.* (2010) stress that rural areas, unlike big cities which lend themselves to easy data gathering, require rigorous sampling by the inevitably smaller groups of volunteers to achieve representative and reliable spatial data for OSM. Foody *et al.* (2015) comment on the difficulty of identifying and favouring one contribution against the majority view provided by other contributors: an accurate label provided by one volunteer out of a million can be lost within the much larger sea of alternative categorization, as consensus-based initiatives always follow majority dominant views. There is a possibility that increasing the number of contributors may degrade rather than enhance the quality of contributed data sets: some studies (Snow *et al.* 2008, Haklay *et al.* 2010, Foody *et al.* 2013) suggest there is a natural limit to the number of valid cases to reach the ‘truth’. Foody *et al.*’s (2013) multiple, ‘crowdsourced’ investigation of the impact of sample size on attribute accuracy of VGI in land cover mapping

engaged 65 volunteers on a classification activity of African forests. Only 7 volunteers successfully classified at least 90% of the 299 cases requested: the other 58 volunteers were disregarded. Validation involved ground truthing of the VGI classification by three experts who determined that, as 'satisfactory' results were obtained, the small number of seven independent volunteers was adequate for establishing the quality of VGI in a land cover mapping activity. Snow *et al.* (2008) confirmed that high accuracies can be achieved from a small number of contributors in a crowdsourcing activity.

Haklay *et al.* (2010) emphasize that there are no observable correlations between the number of contributors and VGI quality (here, attribute accuracy), once they exceed 13. However, Goodchild and Li (2012) argue that Linus' Law is not as effective for geographic facts as it is for text-based projects such as Wikipedia. They argue that using many people alone is not sufficient to characterize trends in geospatial data error.

Positional accuracy can also be determined using the 'many eyes' principle: Haklay *et al.* (2010), consider that the presence of information from more people can actually lead to clutter, suggesting that the first five contributions made to a feature have the most influence (producing statistically significant positional accuracy correctness). Similarly, Basiouka (2010) concluded that there was no clear pattern of improved positional accuracy when the number of contributors increased above five, in her creation of dynamic maps for navigational activities in OSM. The results obtained from her study were deemed 'sufficient' for assessing the positional accuracy of geospatial data. Spielman (2014) summarizes by arguing that collectively generated mapping from several participants, whilst not necessarily accurate, can produce credible maps that are beneficial to many users on a FFP basis.

## 2.5. VGI application in land administration

Applying VGI specifically to improve tenure security and delivery of services in informal and customary settlements in developing countries is increasingly important. Its participatory nature has improved the awareness and confidence of the public in land administration activities. Current VGI initiatives are used to secure land tenure in rural villages of Kenya, Ethiopia, Ghana, Rwanda, Tanzania and Rwanda (Bennett and Alemie 2015, Rahmatizadeh *et al.* 2016, Asiama *et al.* 2017, Siriba and Dalyot 2017). Such participatory initiatives can be considered as improving the lives of citizens in rural areas, especially women and children who tend to be side-lined in traditional systems (Quan and Payne 2008).

VGI has been proposed as a practical and low-cost method for fast acquisition of land information to identify and map land rights, restrictions and responsibilities of communities in developing countries (Rahmatizadeh *et al.* 2016). Further studies (Johnson and Sieber 2013, Zevenbergen *et al.* 2013) have investigated and outlined how VGI could be formalized in official systems for informed decision-making and recognition of social tenures common in rural areas. These studies are significant in identifying and outlining how VGI initiatives adopted in official systems can improve lives of local communities. However, issues of how trust and confidence can be placed on VGI in such applications need further investigation. Where ground truth is non-existent or inadequate, the use of TRM may be of value.

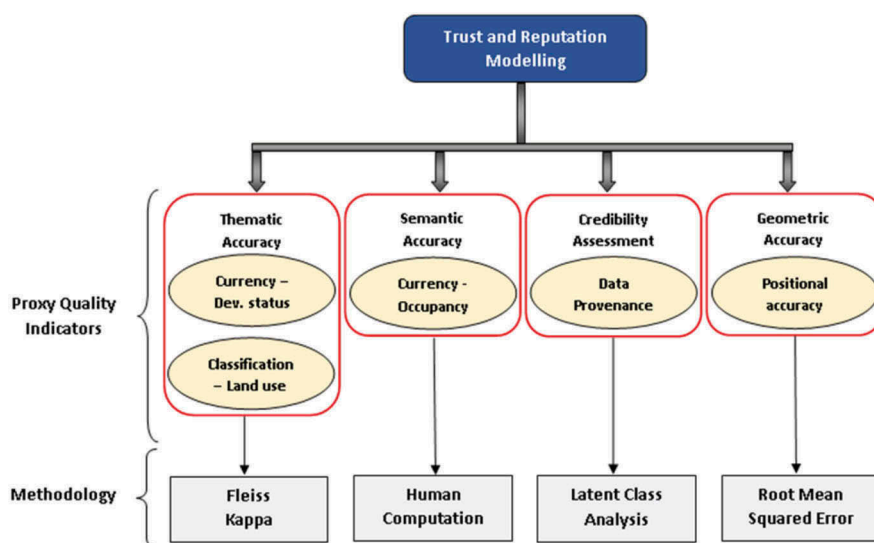
### 3. Methodology for TRM for land administration

#### 3.1. Introducing methods of TRM

This study concentrates on land administration, which involves significant human activity addressing many different tasks, including extensive geospatial data handling related to land parcels. Whilst recognizing that geospatial data collected by the public may potentially contain many errors, TRM is proposed as a method to establish the 'proxy' quality and credibility of VGI produced by volunteers. It is suggested here, in the context of VGI and TRM, that there are four data quality indicators which can contribute to the quantification of trust and reputation: (a) thematic accuracy, (b) semantic accuracy, (c) credibility assessment and (d) geometric accuracy measures (Figure 1). These data quality measures and their methodologies will be elaborated in Section 3.2.

Figure 1 shows the four data quality indicators identified, together with the land parcel parameters to be investigated by each indicator, and the methodologies to be used to establish the quality of VGI and credibility of volunteers. TRM can adopt the 'social', 'crowdsourcing' and 'geographic' approaches introduced by Goodchild and Li (2012), to establish robust quality measures of VGI. These three approaches have been adapted here to a more general TRM framework, such that comprehensive 'proxy' quality and credibility measures of VGI are obtained.

The social approach involves utilizing trusted individuals as gatekeepers to assess and monitor contributions by other volunteers. For example, TRM can use experts (trusted intermediaries) in local government to examine and assess volunteer contributions to establish the quality of their contributions. The crowdsourcing approach assumes that when more people work on the same area, they are likely to identify and correct errors in the data, thus increasing the quality of the data sets. For example, TRM stresses that the more participants agree to a contribution by a volunteer, largely through inaction of



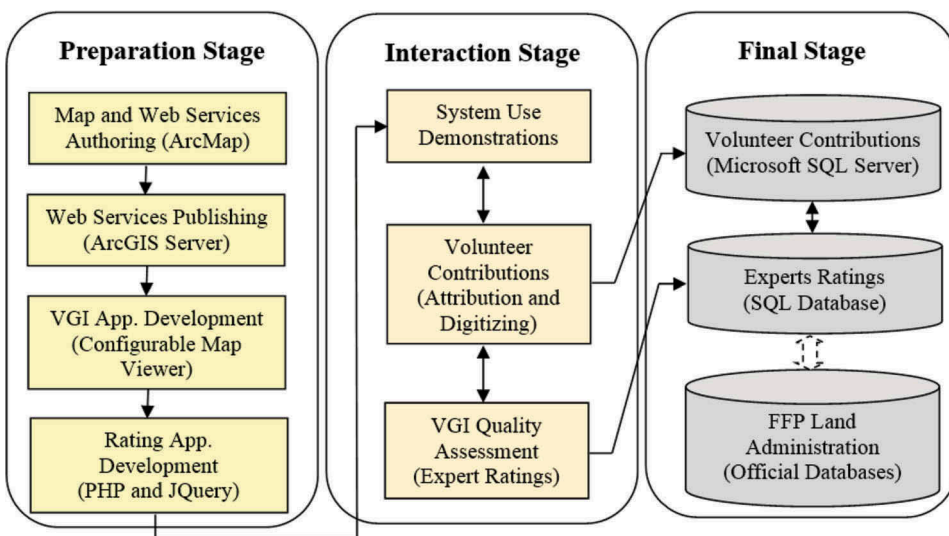
**Figure 1.** Trust and reputation modelling framework for determining proxy VGI quality and credibility.

not editing it or provision of positive reviews about it, the more it can be regarded as reliable and of acceptable quality to other consumers. The geographic approach exploits knowledge from geography to detect unlikely or impossible configurations in the contributed data sets (Ballatore and Zipf 2015). For example, there is general geographic knowledge that land parcels cannot be in the middle of water courses or road networks. Quantifying VGI quality outlines its usefulness in terms of reliability, credibility and potential for incorporation in official systems.

To assess VGI using TRM, a Web map application was developed using ArcGIS JavaScript Application Programming Interface (API). The objective of its development was to provide a test bed for the public to examine a sample of existing geospatial data in official databases, to comment on the currency and accuracy of its contents, and later contribute their own land information of the study area. Prior to its development, a preliminary data collection study using interviews and questionnaires was conducted in the study area engaging with key stakeholders, including community leaders, the public (comprising land owners, aspiring land owners, and tenants) and officials at the governmental land authority, responsible for official land administration.

Semi-structured interviews were used to gather all stakeholder opinions. The interviews were tape recorded and later transcribed and analysed using a discourse analysis methodology. The data collection study provided valuable insights about perceptions of the public and those in the public service regarding the current land administration in the region. The information collected was later used to guide the development of the Web map application and address some of the issues raised by the public. The structure of the application, its components and data flows are presented in Figure 2.

The VGI application framework consists of three main stages of information flow: (a) preparation stage, (b) interaction stage and (c) final stage. Geographical data and Web services were built and published in the *preparation stage* such that they could be consumed by volunteers and experts. To allow volunteers to contribute geospatial data



**Figure 2.** VGI application framework, highlighting the stages of information flow.

sets, a Web map application was further developed in this stage with a Configurable Map Viewer (CMV). A rating application was also developed in the preparation stage with PHP, MySQL and JQuery to provide a platform for experts to rate information contributed by volunteers. The rating application facilitated subjective measures of quality and credibility of contributed data sets for conclusions to be drawn based on consensus agreement. The provision of a platform for experts to assess public contributions enforced the 'social' and 'crowdsourcing' approaches of VGI accuracy assessment (Goodchild and Li 2012).

Initial demonstration of system use was performed in the *interaction stage* and volunteers later engaged in contributing land information in the form of attribute data to predefined land parcels, and digitizing land parcels of different land uses within the study area. Observations of system use and data handling capabilities were conducted when volunteers interacted with the application to conduct a usability evaluation measure of functionalities and user experience. In addition, experts were engaged in this stage to subjectively rate and assess volunteer contributions.

The *final stage* of the application is concerned with the storage of contributed data sets and expert ratings as well as the possible consideration of VGI in official systems. An Enterprise geodatabase based on Microsoft SQL Server and a MySQL database were used to store contributed data sets and expert ratings, respectively. An investigation of the legal framework related to land administration and the practice of the law in land transactions, parcel definitions, occupancy, land use and restrictions etc. provided information on the extent to which data sets produced in the VGI application could be incorporated in official systems for FFP land administration.

### 3.2. Measuring VGI quality using TRM

The four data quality indicators of VGI in TRM (Figure 1) address five specific parameters about contributed data sets: (a) currency evaluation of a land parcel's development status; (b) land use classification accuracy; (c) identification and description of a land parcel's occupancy; (d) data provenance and (e) positional accuracy determination. These parameters are fundamental to the establishment of effective land administration, especially for policy formulation and monitoring components of a LAS (Dale and McLaren 1999, Ayten and Cay 2014). Thematic accuracy of contributed data sets was determined by the first two parameters, semantic accuracy and volunteer credibility by the third and fourth parameters, and geometric accuracy by the final parameter.

#### 3.2.1. Thematic accuracy measure

Thematic accuracy measure examines how well individuals identify and classify objects in the VGI application. The Fleiss Kappa index was used to assess reliability of data contributed by volunteers, quantifying accuracies of land parcel classification by public contributions relative to the experts' decisions. Fleiss Kappa is a generalization of Scott's pi statistic, a statistical measure of inter-rater reliability (Banerjee *et al.* 1999). Unlike Cohen's Kappa and Scott's pi, which work for two raters, Fleiss Kappa is applicable for many raters to process categorical ratings for a fixed number of items: it is the most widely used index of inter-rater agreement for variables with nominal categories (Randolph 2005).



The advantage of using Kappa, compared to other statistical measures of agreement, is that it considers agreement that would be expected by chance. Therefore, it is a more robust measure of agreement than simple percent agreement calculation (Sim and Wright 2005). Nonetheless, a Kappa measure can often provide low values despite high levels of agreement between raters, because of chance correction computation. In addition, a single value of Kappa is difficult to interpret, especially when trying to diagnose the possible cause of a lack of agreement (Byrt *et al.* 1993). It is recommended that a confidence interval be constructed around the obtained value of Kappa to reflect sampling error (Mashour *et al.* 2010, Olofsson *et al.* 2014), providing meaningful interpretations of the minimum and maximum possible values for Kappa: a small range between the lower and upper limit for Kappa depicts a precise and high likelihood of agreement, a large range more imprecision and less likelihood of agreement.

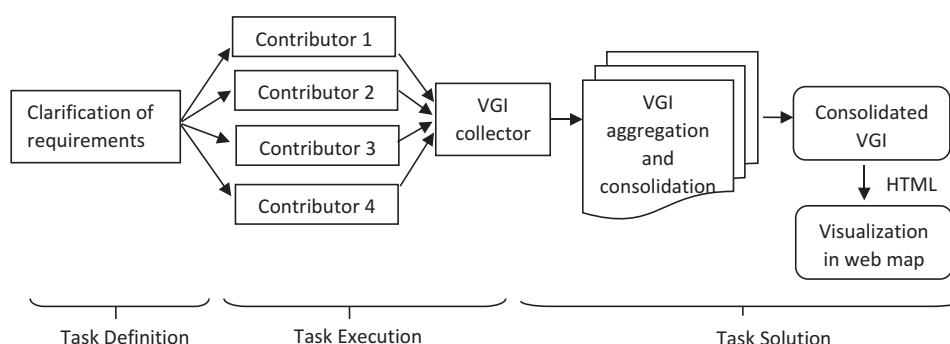
Despite its widespread use, Kappa is discouraged by some: Pontius and Millones (2011) argue that Kappa does not serve a useful role in accuracy measures or area estimations, and Olofsson *et al.* (2014) stress that the corrections of chance agreement produce measures not descriptive of the map accuracies to be encountered by a user. However, in the absence of ground truth, Kappa can be regarded as a reliable tool for 'proxy' accuracy assessment of VGI based on the wisdom of the crowd principle. If many people agree with participants' contributions, this provides valuable information on the reliability of the contributed data, and consensus agreement increases the confidence of potential users of VGI. Kappa, in these instances, is used as a confidence and reliability measure of volunteer contributions by raters, providing quantified agreement between two or more raters who make independent ratings about attributes of a land parcel contributed by volunteers.

### 3.2.2. Semantic accuracy measure

The lack of ground truth in developing countries to help establish the quality and credibility of VGI has motivated the investigation of data provenance as an indicator of VGI quality. Data provenance can provide a valuable dimension when multiple records of the same entity are aggregated to define a final label and improve semantic accuracy. Metadata created in open labelling systems for collaborative projects like VGI can be examined with semantic accuracy measures to collect folksonomies. While a flexible collaborative approach of VGI allows for rich description of geospatial objects to capture local meanings, it also creates semantic heterogeneities: there may be diverse and conflicting attributes used to describe the same object contributed by many users.

Semantic heterogeneity was addressed here by Human Computation (HC) methods (Ballatore *et al.* 2013, Celino 2013, Ronzhin 2015), a technique whereby some computational processes are 'outsourced' to humans. In HC, a computer asks a person or group of people to solve a problem, then collects, interprets and integrates their solutions: in VGI this can consolidate contributed data sets from a variety of sources (Law and Von Ahn 2011, Celino 2013), and addresses the shortcomings of heterogeneous information collection and semantic accuracy challenges common in VGI.

The structure of HC is made up of three steps (Figure 3), (Celino 2013): (1) Task definition, where contribution tasks and requirements are clarified to participants; (2) Task execution, where multiple participants are given similar tasks to contribute information and (3) Task solution, where individual contributions are consolidated and



**Figure 3.** The Human Computation VGI collection and consolidation workflow (Celino 2013).

harmonized into a central solution. HC addresses the semantic heterogeneity of VGI by consolidating similar contributions into single labels, thus improving quality.

Semantic accuracy was computed using Datalift, an open platform for publishing and interlinking data sets on the Web. A semantic query language for databases, SPARQL, was used in Datalift to query, retrieve and manipulate data contributed by volunteers. SPARQL, a recursive acronym for SPARQL Protocol and Resource Description Framework (RDF) Query Language, is a World Wide Web Consortium (W3C) specification (since 2008) used here to aggregate and consolidate VGI based on the tags contributors provide for the same land parcel. The merging process is facilitated by a Natural Language Processing (NLP) technique that automatically combines similar text snippets from multiple sources to form a summary (Manning and Schutze 1999, Barzilay 2003).

SPARQL uses an aggregation algorithm to consolidate VGI tags based on a simple agreement mechanism. Its functionality is such that, as soon as two contributions with similar content from two different volunteers are recorded, the algorithm is triggered and the contributions consolidated into a single occupancy label. Every time a new contribution is made, the algorithm compares it with previously stored labels to determine if consolidation must occur or not. The aggregated results are then displayed as the final label (in the case of a land parcel, this could describe 'occupancy', or 'land use') as HTML.

A RDF – a W3C specification since 2004 – enables source data to be converted into a set of triples (subject, predicate and object) in the task solution stage, for ease of query and integration with other external data over the Web. Here, the subject is the unique identifier (parcel number) of the contributed entity, the predicate is the attribute of the entity (e.g. land use) and the object represents its attribute name (e.g. commercial). The HC approach shows how VGI provenance can be leveraged in a data aggregation and consolidation activity to improve VGI quality based on similar words that volunteers use to describe land parcels in the study area.

### 3.2.3. Volunteer credibility determination

No single measure exists for establishing the credibility of VGI volunteers: this is another challenge in assessing VGI quality and a hindrance for its consideration in authoritative systems. If information is attributable to a known source, it is likely to be trusted more, and have higher reliability and quality, than from the majority of VGI, mostly produced



by unknown volunteers. According to Antoniou and Skopeliti (2015), the emphasis of VGI quality determination has been on the characterization of contributed geospatial data sets, with less emphasis on volunteer credibility: there is a need to investigate how such credibility can be measured based on the content produced.

Statistical methods can be used to analyse and model the relationship between volunteers and their contributions. To establish the credibility of volunteers, a Latent Class Analysis (LCA) methodology is proposed. LCA is recognized as an effective methodology to analyse trends and qualities of multiple contributions from volunteers (Huang and Bandeen-Roche 2004), and has been widely used to assess the accuracy of volunteers in land cover maps (Foody and Boyd 2012, Foody *et al.* 2013). It takes observed variables provided by volunteers to compute information on the unobserved (latent) variable, here representing volunteer reputation. Moreover, LCA can be used to evaluate diagnostic tests without reference to validation by ground truth. The contributions from volunteers were compared against consensus-based classification values of trusted intermediaries via cross-tabulation, to represent final land parcel tags used as input for volunteer reputation computation in Mplus Statistical Analysis software (Muthén 2004, Jung and Wickrama 2008). LCA requires that each observed entity be statistically independent of other variables. Foody *et al.* (2013) used latent class models to measure the accuracy of four volunteers in labelling tropical forests in a 'Globcover' map in West Africa, extracting information on the quality of contributed data sets to establish contributor accuracies in the map without reference to ground truth.

A standard latent class model can be constructed based on the probability of observing patterns of class allocations by a series of classifiers applied to a data set (Foody 2012). These class allocations are known as observed variables (here, land parcel classifications), and are used to provide information on the unobserved variable (which equates to the volunteer reputation). Volunteer reputation has been established using Bayes theorem, which describes the probability of an event happening based on prior knowledge of conditions related to it (Vermunt and Magidson 2003). For example, a person's ability to correctly identify and classify several land parcel parameters can be used to represent a reputation category to which they belong. Therefore, a volunteer's reputation derived from Bayes theorem was allocated to the class which displayed the highest posterior probability of class membership (Vermunt and Magidson 2003, Foody 2012).

In this study, LCA was used to estimate the reputation of volunteers based on their multiple classifications (land use, occupancy and development status) of different land parcels. To achieve this, experts were engaged to assess and rate how well volunteers correctly classified land parcels in the study area. Volunteers with good reputations are characterized by producing geospatial data of good quality. An advantage of LCA is that it can be used to characterize the accuracy of each contributor's labelling regardless of the number of contributions made (Foody *et al.* 2015).

One of the main issues with LCA is determining the number of classes and statistically assessing the fit of each class to the data to obtain representative results. A four-class model was selected in this case to compute volunteer reputations, since it provided a good understanding of reputation categories that volunteers belonged to. Entropy in LCA is used to examine model fit of how well individuals are assigned to membership

(reputation) classes. An entropy value close to 1 shows good model fit and a clear separation of categories (Nylund *et al.* 2007, Jung and Wickrama 2008).

### 3.2.4. Positional accuracy determination

Positional accuracy describes the extent to which a geospatial entity deviates in space from ground truth. A common scientific measure of positional accuracy is the Root Mean Square Error (RMSE) (FGDC 1998). This was computed to determine the amount of deviation between contributed data sets by the participants to those obtained from a Real Time Kinematic (RTK) survey process. RTK survey data was collected with a GPS unit to mimic official data sets of the study area. A RMSE computation was conducted against digitized data sets by the public and those by experts, to determine if expert data could be used as ground truth in the absence of RTK data.

This case study area (Section 4) has applied Botswana Surveying and Mapping Standards (BSMS) for digital geospatial data, endorsing the National Standard for Spatial Data Accuracy (NSSDA) procedure (Ryden 2006). The standards provide an acceptance criterion that geospatial data must conform to, such that it can be considered acceptable for certain uses.

Currently, the study site has digital orthophoto coverage of map scale 1:5000. Table 1 shows horizontal accuracies for digital planimetric data and their associated thresholds as outlined by BSMS. The horizontal accuracy thresholds (Classes I, II, III) provide map accuracy class ranges that contributed data sets can fall into based on the extent of their deviation from ground truth. These classes inform the potential uses of contributed data sets in official systems.

NSSDA recommends a test of a minimum 20 check points per 500 square kilometres, to reflect the geographic area of interest and the distribution error in a data set (FGDC 1998). When 20 check points are tested, a good estimate of the unknown parameter can be attained at 95% confidence level (FGDC 1998). A comparison of the data sets (VGI, experts and RTK data) could provide insights on the positional accuracy levels of contributions by the public. Such information together with the RMSE results can inform the FFP of VGI data sets, with regards to acceptable positional accuracy levels. RMSE is insensitive to outliers (Chai and Draxler 2014) therefore it was difficult to identify individual measurements responsible for uncertainties such that their spatial distributions could be investigated.

## 4. Case study

### 4.1. Applying the principles to Mochudi, Botswana

To examine more fully the role of community mapping and data collection in the Botswana context, the village of Mochudi, one of the nation's larger villages (population 44,815 in 2011) was chosen as a case study area for this research. The village has a

**Table 1.** Horizontal accuracy/quality for planimetric data (1:5000) (GEOMANUAL 2014).

Map scale	Approximate source imagery (digital orthophoto) Ground Sample Distance (GSD)	Horizontal data accuracy class	RMSE <sub>x</sub> or RMSE <sub>y</sub> (cm)	Resultant RMSE (cm)	Horizontal accuracy at the 95% confidence level (cm)
1:5000	40–100 cm	I	62.5	88.4	153.0
		II	125.0	176.8	306.0
		III	187.5	265.2	458.9

mixture of both traditional and modern land development patterns and variable land uses which include residential, civic and community, commercial, industrial and agriculture (commercial and subsistence farming) categories. Botswana has three main land tenure systems: (a) state land (20%); (b) freehold (5%) and (c) customary land (75%): land parcels in Mochudi fall under the latter, which despite being the largest proportion is the most poorly mapped and under-resourced tenure system.

The key stakeholders in this study are the national mapping agency, Department of Surveys and Mapping (DSM), District Land Boards, local community leaders, members of the public and staff at the Land Board. DSM is the main geospatial data source for Land Boards, which are the administrators of customary land. Volunteers were engaged in this study to represent a sample of the local community. They were mandated with contributing land information of the study area using the VGI application developed (Figure 4). Junior staff personnel of the Land Board were engaged as trusted intermediaries (experts) to assess and rate data contributed by the public. Sample spatial data and its corresponding attributes were obtained from the Land Board and their database contents examined and preprocessed. To implement a FFP LAS, Enemark *et al.* (2014) recommend the use of aerial photographs rather than informal field surveys. Thus, 25 cm high-resolution orthophotos were also obtained from the Land Board and used in the VGI application.

A total of 90 members of the public and 15 experts took part in the data collection activity for a period of 6 weeks during the southern hemisphere spring. The objective was for each member of the public to tag and classify 30 predefined land parcels and further digitize 12 land parcels of different land uses in the study area. The averaging of coordinates from multiple contributors was used to define a single vertex (FGDC 1998). Experts were then engaged to inspect, rate and comment on the records contributed by the public.

A total of six zones were created: 30 land parcels of different land uses (industrial, commercial, residential, civic and community, parastatal/government department and agriculture) in each zone were predefined and highlighted with red push pins for volunteers to identify and classify. Each land use category had five different land parcels. Approximately 15 participants were engaged in each zone to tag and label the 30 land

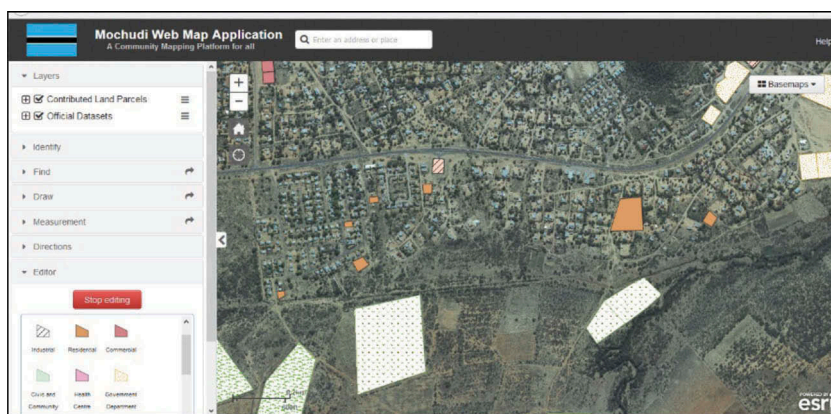


Figure 4. VGI Web map application.

parcels in it, from the orthophoto provided as the visual layer in the VGI application. Therefore, each land parcel was tagged approximately 15 times by participants to obtain multiple records. The total number of contributions using the tagging process were 15 volunteers  $\times$  30 land parcels  $\times$  6 blocks = 2700 records. The objective of the tagging process was to measure how well participants could identify and correctly classify land parcels in their local community.

The results obtained from the consensus agreement of experts about public contributions were then used to measure the 'proxy' quality of contributed data sets and credibility of the public (volunteers) using the methodologies outlined in Figure 1. Another activity by the public involved digitizing predefined land parcel boundaries in the study area. The objective of this activity was to examine how well participants could semantically identify and accurately digitize land parcels.

## 4.2. Assessment and evaluation of data collection results

### 4.2.1. Thematic accuracy results

Geographic information has individual qualities that require modelling to capture agreements between users (Goodchild 2009). For *currency determination* of contributed data sets, experts assessed and rated the *development status* parameters as classified by participants. Fleiss Kappa was then used to measure agreement between the expert rating and the volunteer contributions (derived by majority interpretation).

The degree of agreement of experts who assessed citizen's interpretation of residential land parcel's development status was 82%: thus the vast majority of experts agreed that the land parcels investigated had been correctly classified by the participants. A Kappa statistic measure of 0.65 (0.36–0.94, 95% confidence interval) was obtained (Table 2). This represents a 'substantial agreement' of experts in the public's decisions (Landis and Koch 1977). The Kappa value obtained shows an above average degree of agreement between experts, but with low precision (large confidence interval margin): confidence and trust can be placed on the development status classifications of contributed data sets but caution should be observed.

Adverse outcomes and variations were observed between expert ratings in the assessment of development status classifications of agricultural land parcels. Many classified as 'developed' by the public were regarded by some experts as incorrect. Further investigation revealed that there were no clear distinctions in official systems of what constituted a developed agricultural land parcel and what did not: for example, a land parcel cleared of vegetation and with a boundary fence was classified as 'developed' by the participants, while 45% of the experts considered it 'undeveloped' due to a

**Table 2.** Thematic accuracy measures of contributed data sets.

Parameter	Kappa measures for each block						Average Kappa	Confidence interval margin	Agreement level
	Blk 1	Blk 2	Blk 3	Blk 4	Blk 5	Blk 6			
Dev. status	0.518	0.607	0.572	0.844	0.718	0.653	0.650	0.36–0.94	Substantial
Land use	0.703	0.813	0.820	0.750	0.796	0.795	0.780	0.69–0.87	Substantial

lack of a habitable structure. This is a sign of a lack of clarity in official systems of the definition of appropriate levels of alteration in an agricultural entity. This is not a challenge to VGI when adopted on a FFP basis, because emphasis is on purpose rather than conformance to strict standards-based processes. For example, one of the key principles of the FFP approach states that land information accuracy should be understood as a *relative* issue, concerned with the use of the information to control land use and increase security of tenure, rather than following advanced technical standards (Enemark *et al.* 2014). Other land uses investigated showed minor variations among expert ratings with the modal rating values considered final classifications of the entities.

TRM relies on the ‘power of the crowd’ principle, and in this case land parcels whose land use classes had been rated by at least six experts were considered for the *land use classification measure*. Six different land use types were identified in the study area to classify each land parcel. A six point rating scale related to the classes was designed to help experts provide subjective ratings based on a land use weighted matrix table (Wang *et al.* 2013): rating values of four and above represented values denoting positive accuracy measures of a land use classification, while those of three and below denoted negative accuracy measures.

Out of the 180 (30 in each of six zones) land parcels considered for land use classification accuracy determination, 165 records had been completely rated by experts. Of these, 138 had average ratings of four and above: 84% of the rated land parcels contributed by the public. Nineteen land parcels received average rating scores of four, 69 got a rating of five and 50 land parcels received maximum average rating scores of six. The remaining 27 land parcels received average rating scores of less than three denoting incorrect classifications.

To measure the validity of the results from the frequency distribution calculation, an inter-rater agreement of expert ratings was computed. Land use classes accepted as the final classifications from expert ratings of volunteer inputs were computed to determine the degree to which experts agreed with them. A Fleiss Kappa coefficient value of 0.78 (0.69–0.87, 95% confidence interval) was obtained (Table 2): this is a ‘substantial’ level of agreement, with a small confidence interval range showing a precise and high degree of agreement of experts with land use classifications contributed by the public. This means that confidence and trust can be placed on the reliability of land use classifications by the public in this case.

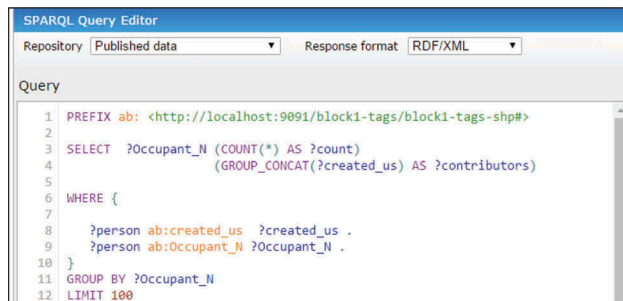
#### 4.2.2. Semantic accuracy results

For the *semantic accuracy measure*, multiple labels from the tagging process were used as input data. This information was then converted into RDF and analysed using SPARQL. The aggregation algorithm of SPARQL was used to merge classifications with similar descriptions (Table 3). On average, a land parcel was classified by 15 volunteers. Figure 5 shows a sample SPARQL query on the provenance data which conducts a count and concatenates all land parcel *occupancy classifications* (‘Occupant\_N’) with similar lexical terms for a single output value.

Table 3 shows a sample of the consolidated output of the HC approach with the SPARQL query results indicating the first-five occupancy parameters, their aggregated outputs and a list of volunteers who provided the consolidated occupancy classifications. From the table, two Supermarkets (Sefalana Hyper and Saverite) have received the

**Table 3.** SPARQL consolidated output of land parcel occupancy and volunteers.

Land parcel occupancy label	Aggregated count	Volunteers (V)
Pilane scrap yard	7	V1, V2, V3, V6, V9, V13, V15
Pilane brick moulding plant	6	V1, V4, V7, V8, V11, V14
Nutri Feeds Botswana	11	V1, V2, V3, V4, V5, V6, V8, V11, V12, V13, V15
Sefalana Hyper Supermarket	12	V1, V3, V4, V5, V7, V8, V9, V10, V11, V13, V14, V15
Saverite Supermarket	12	V1, V2, V3, V4, V5, V7, V8, V9, V10, V12, V13, V14

**Figure 5.** Sample SPARQL query on the provenance data.

highest number of similar classifications (12) from volunteers. The consolidated outputs can later be incorporated into the VGI application as the final occupancy labels of the land parcels. This increases the semantic accuracy of the land parcel's classification, as the occupancy label with the highest number of similar contributions is consolidated as its final classification label.

There were instances where final volunteer occupancy labels differed in large numbers, and this was observed in some commercial enterprises (Table 4). For example, some participants preferred to describe a land parcel's occupancy with the entity's owner name (this is a norm in the area), others used a popular old name, whereas some used its current trading name. Such observations require further investigation by potential stakeholders for the classifications to be trusted and relied upon.

The semantic accuracy measure has shown that semantic heterogeneity of contributed data sets can be addressed through HC methods to establish their trustworthiness and reliability.

#### 4.2.3. Volunteer credibility results

*Volunteer reputation* was established by computing the extent to which volunteers correctly classified land parcels in terms of their land use, occupancy and development status.

**Table 4.** SPARQL consolidated output of different land parcel occupancy labels.

Land parcel occupancy label	Description of label	Aggregated count	Volunteers (V)
Ga Thabo (Thabo's place)	Owner's real name	4	V3, V6, V8, V10
Masakeng Restaurant	Popular old name	4	V1, V4, V5, V13
Phaphane Supermarket	Current trading name	4	V2, V9, V12, V15
Unknown	Unknown	3	V7, V11, V14



**Table 5.** Posterior probabilities of class memberships of volunteers.

Volunteer	Par 1	Par 2	Par 3	Par 4	Par 30	Class A	Class B	Class C	Class D	Final Class
1	1	1	0	1	0	0	0	1	0	C
2	0	1	1	0	1	0	0.997	0	0.003	B
3	1	0	0	1	1	0	0	1	0	C
4	0	1	1	1	0	0	0.013	0.987	0	C
5	1	0	1	1	0	1	0	0	0	A

For each volunteer, Mplus estimates the reputation category they belong to, based on how well they correctly classify land parcel parameters.

The class with the highest probability is chosen to represent the overall reputation of the volunteer. For example, [Table 5](#) shows that volunteer 1 has a single Class C membership, whereas volunteer 2 has two partial memberships of Class B and Class D: his class membership with the highest posterior probability is Class B, so his reputation is allocated there. A labelled snippet of the output file is shown in [Table 5](#), which is a sample of the first 30 land parcels (Par1, Par2, ..., Par30) classified by 5 volunteers, and their partial class memberships (Class A, B, C and D).

The four latent class models have been equated to a Likert scale. A four-class Likert scale of very good ( $75\% \leq r \leq 100\%$ ), good ( $50\% \leq r < 75\%$ ), poor ( $25\% \leq r < 50\%$ ) and very poor ( $0 \leq r < 25\%$ ) was adopted to define reputation categories ( $r$ ).

A scrutiny of the posterior probabilities for the whole of [Table 6](#) shows that Class C has the most volunteers presenting correct classifications, followed by Class A, Class B and lastly Class D. These class categories were allocated reputation categories as follows: Class C – very good, Class A – good, Class B – poor and Class D – very poor. As a percentage, the class category thresholds were awarded values as shown in [Table 6](#). Therefore, a volunteer who belonged to Class C would fall under the ‘very good’ reputation category and within the 75–100% range.

[Table 7](#) shows the overall numbers of volunteers who took part in the classification activity and the reputation categories they belong to. The category with the highest number of volunteers is Class C (very good reputation) at 38.9%. It was followed by Class A, the ‘good’ reputation class which gathered 23 volunteers (25.6%). Combining the two reputation categories (very good and good) provides a total of 58 volunteers out of the total 90 (64.4%). It can be concluded that approximately two-thirds of volunteers were able to correctly classify land parcels in high proportions. This positive sign indicates that volunteers engaged in the data-collection activity have awareness about land information in their local community. Lastly, very few participants, 17 (18.9%) and 15 (16.7%) had ‘poor’ (Class B) and ‘very poor’ (Class D) reputations, respectively. These low numbers are an indication that confidence can be placed on a high proportion of participant’s contributions about their

**Table 6.** Posterior probability classes, their reputation categories and thresholds.

Class	Reputation category	Thresholds (%)
Class C	Very good	75–100
Class A	Good	50–74
Class B	Poor	25–49
Class D	Very poor	0–24

**Table 7.** The overall number of volunteers engaged and their reputations.

Reputation category	Number of volunteers	Percentage of volunteers (%)
Very good	35	38.9
Good	23	25.6
Poor	17	18.9
Very poor	15	16.7
Total	90	100

local community. Content from volunteers with a good record of positive contributions can be trusted by potential consumers on a FFP basis.

#### 4.2.4. Positional accuracy results

Positional horizontal accuracy of contributed land parcels was computed by comparing planimetric coordinates of their well-defined points with coordinates from an independent source of higher accuracy (RTK survey data sets).

To measure the *positional accuracy* of contributed data sets, a RMSE statistic was computed in the horizontal plane. Three assessments were conducted (Table 8): (1) built-up areas (residential); (2) agricultural areas (ploughed fields) and (3) overall RMSE of all data collected. The anticipation was that the three assessments could provide insights about how well participants identified and digitized land parcels with different levels of detail in the orthophoto. Built-up areas are more detailed than agricultural areas, hence expected to produce smaller positional deviations. The tested horizontal accuracy of contributed features in built-up areas against RTK survey data sets was found to be 0.74 m at 95% confidence level. Since a 1:5000 map scale orthophoto was used, this conforms to a horizontal data accuracy Class I of the BSMS accuracy threshold (Table 1): such high accuracy mapping-grade geospatial data can be used for the following purposes (GEOMANUAL 2014): (a) policy formulations; (b) planning; (c) decision-making at village level; (d) land registration and (e) cadastral mapping of sparsely populated areas.

The horizontal positional accuracy determination of agricultural land parcels against RTK survey data sets had a RMSE deviation of 2.07 m at 95% confidence level. Despite not having clearer boundaries, agricultural land parcels produced positional deviations of an acceptable Class II category of the GEOMANUAL (2014) specification. Such land parcels can be used for the following mapping purposes: (a) standard mapping and geographic information systems (GIS) work, (b) general boundary surveys, (c) land registration, (d) reporting of illegal activities, (e) land rights recording of monumental

**Table 8.** RMSE computations for different land uses.

Land use	RMSE for volunteers – RTK (m)	RMSE for volunteers – experts (m)	Horizontal accuracy Class
Built-up areas	0.74	0.66	I
Agricultural areas	2.07	1.73	II
Combined land uses	1.59	1.16	I and II



sites and (f) water points location determination. An overall RMSE error of all contributed data sets (residential, commercial, civic and community, agricultural and industrial) was 1.59 m at 95% confidence level.

The significance of such a result is that confidence and trust can be placed on contributions made by the public regardless of their land uses or spatial location. It implies that when a high resolution orthophoto is used in the VGI application, high accuracy vector data can be produced by on-screen digitizing at an overall acceptable Class II accuracy level.

A further scrutiny of how well experts digitized similar land parcels to volunteer contributions was conducted to establish the amount of deviation between the two data sets. It was observed that the overall RMSE of volunteer contributions against experts' data sets was 1.16 m at 95% confidence level, which was lower than the amount of deviation obtained from RMSE of volunteers against RTK survey data (1.59 m) (Table 7). This result suggests that volunteers can produce land parcels of acceptable Class I and II accuracy levels of geospatial data when compared to digitized data sets of experts. Moreover, these computations have less positional deviations compared with RTK computations, whose data was acquired through high accuracy surveys. Therefore, digitized data sets by experts can be trusted and used as ground truth to measure the positional accuracy of volunteer contributions, if survey data is unavailable.

## 5. Discussions and future work

The outcomes of the TRM methodology are 'proxy' measures of VGI quality and credibility of volunteers. TRM can help people make informed decisions and judgments about the quality, reliability and relevance of information produced by other volunteers without reference to ground truth. It uses the 'wisdom of the crowd' principle and assumes that there are hidden objective truths which can be appropriated from many contributors and their consensus agreements. In this case, the introduction of TRM was motivated by a general lack of access to official data in LAS in developing countries, and the lack of standard accuracy measures of VGI in land administration. Four quality indicators – (a) thematic accuracy, (b) semantic accuracy, (c) volunteer reputation and (d) positional accuracy – were examined to develop the TRM.

Thematic accuracy enforces the trust element of VGI when trusted intermediaries are tasked with an objective of assessing and rating the reliability of volunteer contributions: Kappa statistic measures can be used to determine agreement between trusted intermediaries. A high degree of agreement of expert's ratings has been obtained from volunteer contributions classifying land parcel's development status and land use parameters. A high degree of agreement implies that confidence and trust can be placed in such data sets in any VGI initiative. The second indicator (semantic accuracy) also enforces the trust element indicating that similar content produced by many volunteers can be consolidated into a single value using a W3C standardized language for improved VGI quality, and clutter reduction.

The third indicator (volunteer reputation) enforces the reputation element of TRM, indicating that, using LCA methodology, multiple contributions of an individual can be used to infer the quality and reliability of the data they produce. A good reputation is a sign of trustworthiness, which can be used by participants to establish credibility in

interacting with others in participatory initiatives like VGI. The fourth indicator (positional accuracy) enforces the extent to which digitized data sets produced by volunteers deviate from high accuracy GPS cadastral surveys, or ground truth if it exists. Standards and specifications used in official systems can be used to inform the extent to which data sets from a VGI initiative can add value to official databases. Therefore, the positional accuracy determination informs the FFP of digitized VGI in official databases. TRM can improve confidence and trust of officials and other stakeholders in considering VGI for use in official systems on a FFP basis.

How the TRM results can be communicated in the VGI application for improved quality of volunteer contributions and for informed decision-making needs further investigation. In VGI, a volunteer's reputation value could be placed alongside the attributes of their contributed data sets. In that way, potential consumers can make informed decisions about the data sets before utilizing them. This paper has focused on the potential of the TRM methodology in establishing the quality of contributed data sets and credibility of volunteers, but limitations in land administration should be considered also.

## 6. Conclusions

This article has presented a novel TRM methodology to establish the quality and credibility of VGI such that it can be considered in LAS on a FFP basis, where ground truth is non-existent. All relevant parameters have been assessed: even for positional accuracy requirements, it has been shown that the boundaries of land parcels can be digitized to a Class II standards specification sufficient for identification and determination of extent. Further research is necessary to consider the requirement for ground truth in the specific task of coordinates capture. The overall methodology was tested in a real-world case study in Botswana to present its applicability and potential in official databases. TRM consists of four data quality and volunteer credibility measures (thematic accuracy, semantic accuracy, volunteer credibility and positional accuracy) to ensure that trust and confidence are applied to the use of VGI in LAS without reference to ground truth. Therefore, it provides a possibility by which land information can be collected through the involvement of local communities to improve official databases in developing countries.

VGI has the potential to address inefficiencies in administering land in developing countries, especially rural Africa. For example, TRM has proved that VGI can produce quality and reliable data sets of Class II positional accuracy which can be used to conduct regular systematic updates of geographic information in official systems. Moreover, positional accuracy of volunteer contributions measured against experts' contributions produced reliable data sets of an overall Class I accuracy level. This suggests that trust and confidence can be placed on expert contributions for use as ground truth if official data sets are unavailable or insufficient.

TRM parameters have further validated the significance of the 'power of the crowd' principle such that VGI can provide valuable information for LAS to efficiently handle land changes highly relevant to the community they serve. In a local community, metadata about contributed data sets can be created by the development of rating applications for the public to police themselves in assessing and subjectively rating the

accuracy of other volunteer contributions. Fleiss Kappa can then be used to establish 'proxy' quality measure of VGI based on the extent to which assessors agree on the accuracy of contributions. TRM through HC methods has demonstrated that the semantic heterogeneity common in VGI can improve the quality of contributed data sets by aggregating contents with similar lexical terms. This can increase the prospect of adoption of VGI into official databases, as HC argues that a correct classification of an entity can be obtained from an aggregation of contributions with similar vocabulary.

VGI quality assessment in this study is taken as a citizen science initiative to harness volunteer skills to execute tasks of contributing land information and to promote reputable volunteers to trusted intermediary status. Moreover, quantifying data quality and credibility in VGI underpins its usefulness in terms of reliability and trustworthiness. It further informs its potential for incorporation into official systems. The lack of ground truth in developing countries should not be a hindrance for investigating the possibility of VGI adding value and flexibility to official systems. The TRM methodology proposed here is not suggested as a replacement for conventional and rigorous accuracy measures, but as an alternative means of providing valuable land information to LAS in developing countries. This method has been applied to successfully demonstrate that VGI can produce accurate and reliable data sets which can be used to conduct regular systematic updates of geographic information in official systems.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Kealeboga K. Moreri  <http://orcid.org/0000-0002-3692-1915>

## References

- Adler, B. and de Alfaro, L., 2007. A content-driven reputation system for the Wikipedia. In: C. Williamson, et al., eds. *16th international conference on World Wide Web (WWW)*. New York, NY, USA: ACM Press.
- Ali, A. and Schmid, F., 2014. Data quality assurance for volunteered geographic information. In: M. Duckham, et al., eds. *8th International Conference on Geographic Information Science*. Vienna, Austria: Springer International Publishing Switzerland.
- Antoniou, V. and Skopeliti, A., 2015. Measures and indicators of VGI quality: an overview. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2, 345–351. doi:10.5194/isprsannals-II-3-W5-345-2015
- Artz, D. and Gil, Y., 2007. A survey of trust in computer science and the semantic web. *Journal of Web Semantics*, 5, 58–71. doi:10.1016/j.websem.2007.03.002
- Asiama, K., Bennett, R., and Zevenbergen, J., 2017. Participatory land administration on customary lands: a practical VGI experiment in Nanton, Ghana. *International Journal of Geoinformation*, 6, 1–22.
- Ayten, T. and Cay, T., 2014. The effect of land consolidation components on parcellation plans. In: *Proceedings of the XXV FIG International Congress 2014 - Engaging the Challenges: enhancing the Relevance*. Kuala Lumpur, Malaysia: FIG.
- Babbie, E., 2007. *The practice of social research*. Belmont, USA: Thomson Wadsworth.

- Bajari, P. and Hortacsu, A., 2003. The winner's curse, reserve prices, and endogenous entry: empirical insights from eBay auctions. *The RAND Journal of Economics*, 34, 329–355. doi:[10.2307/1593721](https://doi.org/10.2307/1593721)
- Ballatore, A., Bertolotto, M., and Wilson, D., 2013. Computing the semantic similarity of geographic terms using volunteered lexical definitions. *International Journal of Geographical Information Science*, 27, 2099–2118. doi:[10.1080/13658816.2013.790548](https://doi.org/10.1080/13658816.2013.790548)
- Ballatore, A. and Zipf, A., 2015. A Conceptual quality framework for volunteered geographic information. In: F.S. Irina, et al., eds. *COSIT 2015, Conference on Spatial Information Theory XII Santa Fe*. New Mexico, USA: Lecture Notes in Computer Science.
- Banerjee, M., et al., 1999. Beyond Kappa: a review of interrater agreement measures. *The Canadian Journal of Statistics*, 27, 3–23. doi:[10.2307/3315487](https://doi.org/10.2307/3315487)
- Barzilay, R. 2003. *Information Fusion for Multidocument Summarization*. PhD, Columbia University.
- Basiouka, S., 2010. The use of dynamic maps and volunteered geographic information in Greece. In: *Proceedings of the Joint International Federation of Surveyors (FIG) Commission 3 and Commission 7 Workshop - Information and Land Management. A Decade after the Millenium*, Sofia, Bulgaria.
- Bennett, R. and Alemie, B., 2015. Fit-for-purpose land administration: lessons from urban and rural Ethiopia. *Survey Review*, 1, 1–10.
- Biraro, M., Bennett, R., and Lemmen, C., 2015. Accelerated land administration updates. In: J.A. Zevenbergen, W. Vries, and R. Bennett, eds.. *Advances in responsible land administration*. Raton, FL: CRC Press.
- Bishr, M. and Kuhn, W., 2007. Geospatial information bottom up: A matter of trust and semantics. In: S. Fabrikant and M. Wachowicz, eds. *The European information society – leading the way in geo-information*. Verlag Berlin Heidelberg: Springer.
- Bishr, M. and Mantelas, L., 2008. A trust and reputation model for filtering and classifying knowledge about urban growth. *GeoJournal*, 72, 229–237. doi:[10.1007/s10708-008-9182-4](https://doi.org/10.1007/s10708-008-9182-4)
- Byrt, T., Bishop, J., and Carlin, J., 1993. Bias, prevalence and Kappa. *Journal of Clinical Epidemiology*, 46, 423–429. doi:[10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)
- Celino, I., 2013. Human computation VGI provenance: semantic web-based representation and publishing. *IEEE Transactions on Geoscience and Remote Sensing*, 51, 5137–5143. doi:[10.1109/TGRS.2013.2252015](https://doi.org/10.1109/TGRS.2013.2252015)
- Chai, T. and Draxler, R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7, 1247–1250. doi:[10.5194/gmd-7-1247-2014](https://doi.org/10.5194/gmd-7-1247-2014)
- Collins, A. and Mitchell, M., 2017. Revisiting the World Bank's land law reform agenda in Africa: the promise and perils of customary practices. *Journal of Agrarian Change*, 1, 1–20.
- Corcoran, P. and Mooney, P., 2013. Characterising the metric and topological evolution of OpenStreetMap network representations. *The European Physical Journal - Special Topics*, 215, 109–122. doi:[10.1140/epjst/e2013-01718-2](https://doi.org/10.1140/epjst/e2013-01718-2)
- Cowan, T., 2013. *A framework for investigating volunteered geographic information relevance in planning*. Canada: Msc, University of Waterloo.
- D'antonio, F., Fogliaroni, P., and Kauppinen, T., 2014. VGI edit history reveals data trustworthiness and user reputation. In: J. Huerta, S. Schade, C. Granell, eds. *Proceedings of the 17th AGILE conference on geographic information science, connecting a digital Europe through location and space, Lecture Notes in Geoinformation and Cartography*. Spain: Springer.
- Dale, P. and McLaren, R., 1999. GIS in land administration. In: P.A. Longley, et al., eds. *Geographical information systems: management issues and applications*. Wiley: New York.
- Dellarocas, C., 2002. Goodwill hunting: an economically efficient online feedback mechanism for environments with variable product quality. In: J. Padget, et al., eds. *Workshop on Agent mediated electronic commerce IV: designing mechanisms and systems, Lecture Notes in Artificial Intelligence*, Bologna, Italy: Springer.
- Enemark, S., et al., 2014. Fit-for-purpose land administration. *International Federation of Surveyors (FIG)*, 60, 1–39.

- FGDC 1998. Content standard for digital geospatial metadata. In: *METADATA AD HOC WORKING GROUP*, F. G. D. C. S. (ed.). 590 National Center, Reston, Virginia 20192.
- Fairbairn, D. and Al-Bakri, M., 2013. Using geometric properties to evaluate possible integration of authoritative and volunteered geographic information. *International Journal of Geographical Information Science*, 2, 349–370.
- FAO, 2007. Good governance in land tenure and administration. In: R. Grover, et al., eds. *FAO land tenure studies*. Rome, Italy: Food and Agriculture Organization of the United Nations.
- Flanagin, A. and Metzger, M., 2008. The credibility of volunteered geographic information. *GeoJournal*, 72, 137–148. doi:10.1007/s10708-008-9188-y
- Fonte, C., et al., 2015. Good practice guidelines for assessing VGI data quality. In: F. Bacao, M.Y. Santos, and M. Painho, eds. *AGILE 2015 Lecture Notes in Geoinformation and Cartography*, Lisbon.
- Foody, G., 2012. Latent class modeling for site and non-site specific classification accuracy assessment without ground data. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 2827–2838. doi:10.1109/TGRS.2011.2174156
- Foody, G. and Boyd, D., 2012. Exploring the potential role of volunteers citizen sensors in land cover map accuracy assessment. In: T. Tenbrink, et al., eds. *10th international symposium on spatial accuracy assessment in natural resources and environmental sciences*. Brazil: Florianopolis-SC.
- Foody, G., et al., 2013. Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project: accuracy of VGI. *Transactions in GIS*, 17, 847–860. doi:10.1111/tgis.2013.17.issue-6
- Foody, G., et al., 2015. Accurate attribute mapping from volunteered geographic information: issues of volunteer quantity and quality. *The Cartographic Journal*, 52, 336–344. doi:10.1179/1743277413Y.0000000070
- GEOMANUAL, 2014. *Department of surveys and mapping geomatics manual*. Gaborone, Botswana: Department of Surveys and Mapping.
- Golbeck, J., 2008. Weaving a Web of Trust. *Journal of Computer Science*, 321, 1640–1641.
- Golder, S. and Huberman, B., 2006. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32, 198–208. doi:10.1177/0165551506062337
- Goodchild, M., 2009. The quality of geospatial context. In: K. Rothermel, et al., eds. *Quality of context: first international workshop, QuaCon 2009*. Stuttgart, Germany: Berlin: Springer.
- Goodchild, M. and Li, L., 2012. Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1, 110–120. doi:10.1016/j.spasta.2012.03.002
- Hagenauer, J. and Helbich, M., 2012. Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. *International Journal of Geographical Information Science*, 26, 963–982. doi:10.1080/13658816.2011.619501
- Haklay, M., et al., 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, 315–322. doi:10.1179/000870410X12911304958827
- Haklay, M., Singleto, A., and Parker, C., 2008. Web Mapping 2.0: the neogeography of the geoweb. *Geography Compass*, 3, 2011–2039. doi:10.1111/j.1749-8198.2008.00167.x
- Heipke, C., 2010. Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 550–557. doi:10.1016/j.isprsjprs.2010.06.005
- Huang, G. and Bandeen-Roche, K., 2004. Building an identifiable latent class model with covariate effects on underlying and measured variables. *Psychometrika*, 69, 5–32. doi:10.1007/BF02295837
- Javanmardi, S., Lopes, C., and Baldi, P., 2010. Modeling user reputation in wikis. *Statistical Analysis and Data Mining*, 3, 126–139.
- Johnson, P. and Sieber, R., 2013. Situating the Adoption of VGI by Government. In: D. Sui, S. Elwood, and M. Goodchild, eds. *Crowdsourcing geographic knowledge: volunteered Geographic Information (VGI) in Theory and Practice*. Dordrecht Heidelberg New York London: Springer.

- Josang, A. and Golbeck, J., 2009. Challenges for robust trust and reputation systems. In: C. Jensen, et al, eds. *5th International Workshop on Security and Trust Management (STM 2009)*, Saint Malo, France: Springer.
- Jung, T. and Wickrama, K., 2008. An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2, 302–317. doi:[10.1111/j.1751-9004.2007.00054.x](https://doi.org/10.1111/j.1751-9004.2007.00054.x)
- Kessler, C. and De Groot, R., 2013. Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. In: D. Vandenbroucke, B. Bucher, J. Crompvoets, eds. *16th AGILE Conference on Geographic Information Science*. Leuven, Belgium: Springer Lecture Notes in Geoinformation and Cartography 2013.
- Kwan, M. and Ramachandran, D., 2009. Trust in ONLINE REPUTATION SYSTEMS. In: J. Golbeck, ed. *Computing with social trust. human computer interaction series*. London, UK.: Springer-Verlag London Limited.
- Landis, J. and Koch, G., 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174. doi:[10.2307/2529310](https://doi.org/10.2307/2529310)
- Law, E. and Von Ahn, L., 2011. *Human computation*. Vermont: Morgan and Claypool Publishers.
- Manning, C. and Schütze, H., 1999. *Foundations of statistical natural language processing*. Cambridge, Mass: London.
- Mashour, G., et al., 2010. A novel classification instrument for intraoperative awareness events. *International Anesthesia Research Society*, 110, 813–815.
- Mooney, P., Corcoran, P., and Winstanley, A., 2010. Towards quality metrics of OpenStreetMap. In: D. Agrawal, et al., eds. *SIGSPATIAL international conference on advances in geographic information systems*. San Jose, California, USA: ACM Sigspatial GIS 2010.
- Moreri, K., Fairbairn, D., and James, P., 2015. Technological solutions for citizens' participation into cadastral mapping. In: G. Gartner, et al., eds. *27th International Cartographic Conference 16th General Assembly - Maps Connecting the World*, Rio de Janeiro, Brazil. International Cartographic Association (ICA).
- Mozhgan, T. 2012. *On Some Challenges for Online Trust and Reputation Systems*. PhD, Norwegian University of Science and Technology.
- Mui, L. 2002. *Computational Models of Trust and Reputation: Agents, Evolutionary Games and Social Networks*. PhD, Massachusetts Institute of Technology.
- Muthén, B., 2004. *Latent variable analysis: growth mixture modeling and related techniques for longitudinal data*. Newbury Park, CA: Sage Publications.
- Nylund, K., Asparouhov, T., and Muthén, B., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 535–569. doi:[10.1080/10705510701575396](https://doi.org/10.1080/10705510701575396)
- Olofsson, P., et al., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42–52.
- Osterman, F. and Spinsanti, L., 2011. *A conceptual workflow for automatically assessing the quality of volunteered geographic information for crisis management proceedings of the AGILE 2011*. Netherlands: Utrecht, 1–6.
- Pontius, R. and Millones, M., 2011. Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32, 4407–4429. doi:[10.1080/01431161.2011.552923](https://doi.org/10.1080/01431161.2011.552923)
- Quan, J. and Payne, G., 2008. Secure land rights for all. In: T. Naudin, ed. *Nairobi*. Kenya: United Nations Human Settlements Programme (UN-HABITAT).
- Rahmatizadeh, S., Rajabifard, A., and Kalantari, M., 2016. A conceptual framework for utilising VGI in land administration. *Land Use Policy*, 56, 81–89. doi:[10.1016/j.landusepol.2016.04.027](https://doi.org/10.1016/j.landusepol.2016.04.027)
- Randolph, J., 2005. Free-Marginal Multirater Kappa (multirater Kfree): an alternative to fleiss' fixed marginal multirater Kappa. In: *Joensuu Learning and Instruction Symposium 2005*, Joensuu, Finland: CiteSeerX.
- Raymond, E., 2001. *The Cathedral and the Bazaar: musings on Linux and Open Source by an Accidental Revolutionary*. rev. Sebastopol, CA: O'Reilly.



- Resnick, P., et al., 2000. Reputation systems. *Communications of the ACM*, 43, 45–48. doi:[10.1145/355112.355122](https://doi.org/10.1145/355112.355122)
- Ronzhin, S. 2015. *Semantic enrichment of Volunteered Geographic Information using Linked Data: a use case scenario for disaster management*. Unpublished MSc Thesis, University of Twente.
- Ryden, A. 2006. Specifications for data acquisition topography base map national mapping DSM Botswana. In: MAPPING, S. R. F. T. D. O. S. A. (ed.). *Towards Botswana's National Spatial Data Infrastructure*. Gaborone, Botswana: Department of Surveys and Mapping.
- Sabater, J. and Sierra, C., 2005. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24, 33–60. doi:[10.1007/s10462-004-0041-5](https://doi.org/10.1007/s10462-004-0041-5)
- Sim, J. and Wright, C., 2005. The Kappa statistic in reliability studies: use interpretation, and sample size requirements. *Journal of American Physical Therapy Association*, 85, 257–268.
- Siriba, D. and Dalyot, S., 2017. Adoption of volunteered geographic information into the formal land administration system in Kenya. *Land Use Policy*, 63, 279–287. doi:[10.1016/j.landusepol.2017.01.036](https://doi.org/10.1016/j.landusepol.2017.01.036)
- Snow, R., et al., 2008. Cheap and fast – but is it good? Evaluating non-expert annotations for natural language tasks. In: M. Lapata, H.T. Ng, S. Pado, eds. *Conference on Empirical Methods in Natural language Processing*, Hawaii, USA. SIGDAT (ACL Special Interest Group).
- Sonja, U., Matzat, U., and Snijders, C., 2009. On-line reputation systems: the effects of feedback comments and reactions on building and rebuilding trust in on-line auctions. *International Journal of Electronic Commerce*, 13, 95–118. doi:[10.2753/JEC1086-4415130304](https://doi.org/10.2753/JEC1086-4415130304)
- Spielman, S., 2014. Spatial collective intelligence? Credibility, accuracy and volunteered geographic information. *Cartography and Geographic Information Systems*, 41, 1115–1124.
- Tavakolifard, M. and Almeroth, K., 2012. A taxonomy to express open challenges in trust and reputation systems. *Journal of Communications*, 7, 538–551. doi:[10.4304/jcm.7.7.538-551](https://doi.org/10.4304/jcm.7.7.538-551)
- Tulloch, D., 2008. Is volunteered geographic information participation? *GeoJournal*, 72, 173–183.
- Van Exel, M., Dias, E., and Fruijtjer, S., 2010. The impact of crowdsourcing on spatial data quality indicators. In: S.I. Fabrikant, et al., eds. *GiScience 2011, Lecture Notes in Computer Science*. Zurich, Switzerland: Springer.
- Vermunt, J. and Magidson, J., 2003. *Latent class analysis*. California, Thousand Oaks.
- Wang, Y., Kockelman, K., and Wang, X., 2013. The impact of weight matrices on parameter estimation and inference: a case study of binary response using land use data. *Transport and Land Use*, 6, 75–85. doi:[10.5198/jtlu.v6i3.351](https://doi.org/10.5198/jtlu.v6i3.351)
- Zacharia, G., Moukas, A., and Maes, P., 1999. Collaborative reputation mechanisms in electronic marketplaces. In: *32nd Hawaii International Conference on System Sciences*. Hawaii, USA: IEEE Xplore Digital Library.
- Zevenbergen, J., 2002. A systems approach to land registration and cadastre. In: *XXII FIG International Congress : session TS7.11 3D Cadastre*. Washington, D.C: USA.
- Zevenbergen, J., et al., 2013. Pro-poor land administration: principles for recording the land rights of the underrepresented. *Land Use Policy*, 31, 595–604. doi:[10.1016/j.landusepol.2012.09.005](https://doi.org/10.1016/j.landusepol.2012.09.005)