Editorial

Highlighting Current Trends in Volunteered Geographic Information

David Jonietz [1],*, Vyron Antonio [2], Linda See [3] and Alexander Zipf [4]

[1] Chair of Geoinformation Engineering, ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland; jonietzd@ethz.ch

[2] Hellenic Military Geographical Service, 4 Evelpidon, 11362 Athens, Greece; v.antoniou@ucl.ac.uk

[3] International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, A-2361 Laxenburg, Austria; see@iiasa.ac.at

[4] GIScience Research Group, Institute of Geography, University of Heidelberg, 69120 Heidelberg, Germany; zipf@uni-heidelberg.de

* Correspondence: jonietzd@ethz.ch; Tel.: +41-44-633-32-77

Abstract: Volunteered Geographic Information (VGI) is a growing area of research. This Special Issue aims to capture the main trends in VGI research based on 16 original papers, and distinguishes between two main areas, i.e., those that deal with the characteristics of VGI and those focused on applications of VGI. The topic of quality assessment and assurance dominates the papers on VGI characteristics, whereas application-oriented work covers three main domains: human behavioral analysis, natural disasters, and land cover/land use mapping. In this Special Issue, therefore, both the challenges and the potentials of VGI are addressed.

Keywords: volunteered geographic information; crowdsourcing; OpenStreetMap; quality

1. Introduction

In the last decade, emerging Web 2.0 technologies have dramatically changed the way informational content is generated on the web. Further reinforced by the mass distribution of modern GNSS-enabled (Global Navigation Satellite System) smart phones, these trends have led to an unprecedented growth in VGI (Volunteered Geographic Information) [1], with big spatial data sets now available from numerous web platforms or Location-based Social Networks (LBSN). Researchers soon recognized both the potential and the challenges of using VGI as a complement, reference, or replacement for more traditional commercial or authoritative data, and have increasingly focused on the technological, methodological, conceptual, and social dimensions of this new data stream. Yet research interest in VGI has not declined, which is indicated by the high number of related academic publications in international journals and conferences, ongoing activities such as COST (Cooperation in Science and Technology) Action networks TD1202 (Mapping and Citizen Sensor) and IC1203 (ENERGIC), alliances such as the European Citizen Science Association (ECSA) and its counterpart in the USA (Citizen Science Association), among many others.

The consistent prevalence of VGI-related research motivated this Special Issue, which called for original papers focusing on all topics involving the collection, processing, analysis, and general use of VGI. After the review process, 16 papers were published, which address a broad range of related issues. The aim of this editorial is to capture the main trends in current VGI research as represented by the contents of our Special Issue. Table 1 lists the published papers, which have been classified thematically to illustrate the main topics covered here. At the highest level we can separate the papers into those that deal with the characteristics of VGI, exhibiting a more data-driven approach, while

the latter are focused on applications of VGI. The inherent characteristics of VGI result mainly from the typical lack of formal specifications, and are addressed in this Special Issue in terms of method development for quality assessment and assurance, as well as a conceptual review of existing VGI approaches. From the range of potential fields of application of VGI, contributors have used VGI to analyze human activities, natural hazards, and for land cover mapping. This editorial is structured in accordance to this thematic classification, and briefly reviews the papers in the following sections.

2. Characteristics of VGI

This section describes the main findings of the papers in this Special Issue which deal with the characteristics of VGI such as the terminology and trends, quality assessment, and methods of quality assurance and protocols.

2.1. Review and Classification of Existing Approaches

In the paper by See et al. [2], many different terms that refer to user-generated content were compiled, in particular those that refer to the collection of spatial data by citizens. These terms were then characterized according to whether they represent the information or the process by which the information is collected and whether the data are actively or passively contributed. The evolution of the different terms was then explored over time using Google Trends and a keyword search of the scientific literature. The second part of the paper attempted to understand the current state of VGI by systematically reviewing around 100 projects using a range of criteria such as thematic area, type of data, the expertise and training required, the availability of the data, the quality and reuse of the data, whether information is collected about the participants, and what incentives are offered for participation. The review clearly indicated areas where further research is needed, e.g., data interoperability, incentives for sustainability, and copyright and ownership of the data, among others.

2.2. Quality Assessment

Data quality represents a major issue and a potential impediment for applications of crowdsourced data. Assessing VGI quality is therefore still of substantial interest to the research community. In this Special Issue, this topic is addressed in three studies. Davidovic et al. [3], for instance, focus on the semantic dimension of VGI quality by analyzing tagging practices in OpenStreetMap (OSM). Based on data from 40 cities around the world, they examined whether contributors follow the tagging guidelines as suggested on the Map Features page, and further analyzed the spatial variations of the results. After selecting a set of 10 suitable, frequently used tags and extracting the information from their corresponding Map Features Page, the authors imported the raw OSM data for the selected cities, and assessed and ranked the compliance for each tag and city on a Likert scale. They found that compliance is generally poor, with strong variations in tag usage and compliance between cities.

In the second paper by Mahabir et al. [4], the spatial scope shifts from global to developing countries. Following the traditional approach of comparing authoritative with crowdsourced data, the authors assessed the quality of road data obtained from OSM and Google Map Maker in Kenya. Focusing particularly on coverage, the authors compared these crowdsourced data to official data, and spatially related the results to population density. First, the coverage was computed for each data set on a cell-by-cell basis, and the pairwise differences were calculated. Then, descriptive statistics were calculated for each cell and data set, and significant spatial patterns were identified based on road coverage and population density. The results showed that, in general, the authoritative data were the most complete; however, OSM and Google Map Maker data showed higher coverage in certain types of areas, e.g., slums, as well as the typical reduction in data quality in rural areas.

Finally, the paper by Touya et al. [5] places an emphasis on the methods available for VGI quality assessment, comparing intrinsic and extrinsic approaches. On the example of OSM POIs (Points of Interest) in the city of Paris, the authors assessed the potential as well as the usability of an approach that combines these two methods in one data set in order to compensate for their respective

shortcomings. In particular, with regard to intrinsic data characteristics, they examined the editing history of features with regard to changes in the name, type, and position, and the plausibility of the spatial as well as topological relationships of the POI with corresponding building footprint polygons. In addition, the OSM data were compared to authoritative reference data by matching corresponding POIs, and assessing their positional and thematic quality. In a further step, with geo-tagged photographs from Flickr, a complementary VGI source was introduced to the quality assessment process, which served to assess positional quality. By combining these methods, the authors concluded that a holistic approach to VGI quality assessment was needed, both to improve their individual results and to gain new information about the data set.

2.3. Quality Assurance and Protocols

Closely related to quality assessment are efforts aimed at providing the means to increase the quality of VGI, either by refining data collection processes through protocols or by improving the processes followed during the lifecycle of a VGI project. To this end, five papers covering this topic have been published in this Special Issue, but each one addresses this challenging issue from a different viewpoint. First, Gómez-Barrón et al. [6] propose a methodology for designing and developing VGI systems. The concept of "system" is central to their work as they consider VGI projects as a special kind of an ICT (Information and Communication Technologies) system where the three main components (i.e., technology, people, and organizations) should collaborate harmonically in order to produce the desired outcome. Hence, the development of a VGI system should follow several phases in which each one plays a role in aligning these three components so as to reduce the friction in the VGI production flow. This process can be refined and improved by using feedback loops. The authors evaluated their approach using two case studies that included several existing VGI projects.

In a research vein, Leibovici et al. [7] examine the need to separate two important processes of the data curation effort, namely, Quality Assurance (QA) and Data Conflation or Data Fusion (DCDF). Through the examination of two different cases, one on land cover validation and another on flood inundation extent estimation, the authors challenged the quality assessment of the end product and highlighted the need for greater attention when defining the role that QA and DCDF can play in the overall improvement of VGI quality. They concluded that in the course of a VGI project, the separation of QA and DCDF should be the goal of any data curation design as this will not only improve the quality of the outcome, but will also enhance the re-usability of VGI in other applications and projects.

While these two previous papers offer a high-level overview of a VGI project's function and suggest methodologies to improve the processes of VGI quality assurance, the next three papers, in contrast, approach VGI quality by considering more low-level issues such as protocols, object classification, and attribution processes in order to improve overall data quality. In this context, Mooney et al. [8] provide a generic protocol for the creation of vector-based VGI content. The authors consider three different cases, namely, manual vectorization from maps and imagery, field survey, and bulk data import, and provide a protocol suitable to be implemented in any VGI project that tries to balance the need for rigorous data collection with the motivation of VGI project participants to follow the protocol as well as the freedom and flexibility that a volunteer-based project should provide to its participants. Through a step-by-step approach, the paper describes the challenges that various stakeholders of a VGI project should consider. Thus, the protocol could be of use for spatial data experts, the VGI project community/initiators, ICT experts, and users/contributors, as it provides best practices that should be followed at each stage of a VGI project.

Narrowing the focus to specific spatial data quality elements, Ali et al. [9] propose a methodology for the classification of OSM features that aims at improving attribution. The authors present a guided classification system that tries to conceptually disambiguate overlapping feature classes of land cover and land use. Such cases require considerable expertise that contributors to a VGI project usually lack. The authors then empirically examined the effectiveness of their guided classification process and demonstrated the applicability of the proposed approach.

Finally, along the same lines, Bordogna et al. [10] outline the development of a fuzzy ontology for improving the attribution process of a VGI project. Using the notion of "contextualized VGI", the paper presents a case study in agriculture where contextualized VGI of crop observations is realized through an application that uses an ontology and geographic context. Using a fuzzy ontology with uncertainty level-based approximate reasoning, the authors define the contributors' uncertainty and imprecise observations. This process enables end users of the VGI project to set their own quality standards according to their needs and exploit the data sets produced through such proposed filtering mechanisms.

3. Applications of VGI

This section provides an overview of the papers that focus on the three main application areas of VGI in this Special Issue: human activities; natural hazards; and land cover mapping.

3.1. Analysis of Human Activities

As two papers in this application area demonstrate, VGI has great potential as a data source for the analysis of human spatial activities. In the first paper by Wang et al. [11], the usefulness of social media data as an alternative to traditional survey-based methods for delimiting trade areas, i.e., regions where the potential customers of individual businesses are located, is explored. For this purpose, data from more than 2.4 million users of the micro-blogging service Sina Weibo were used to delimit trade for selected retail agglomerations in Beijing, China. In a two-step process, first users that had checked in at a particular retail agglomeration were identified, following which all their check-in locations were clustered to reveal their individual activity centers. Based on the data, observed visitation frequency and travel distances were computed, and combined with retail agglomeration attractiveness based on their size, in order to serve as parameters for a Huff model. Different trade area delimitation sets, including subsets of users based on the number of retail agglomerations visited as well as cell-based spatial aggregations of user activity centers, were tested and evaluated for their usefulness. The authors found that the sets obtained by aggregating user activity centers had a better delimiting effect, and were therefore used to calibrate the Huff model for trade area analysis.

The topic of VGI as a potential alternative to traditional survey data is also the focus of Heikinheimo et al. [12], who compared data obtained from Instagram with the results of a large visitor survey of a popular national park in Finland in order to assess its usefulness for visitor monitoring. The data were analyzed to answer questions with regards to popular locations within the park, visitor activities, temporal patterns of visitation and activities, the home locations of visitors, and their motivations for visiting the park, among others. For this, Instagram posts were further analyzed, including their aggregation to larger regions, manual classification of the image contents, and identification of the potential home location based on their post history. Comparing the analysis results to the survey data, the authors found strong similarities, and concluded that social media can indeed complement and enrich traditional survey data in order to derive insights about the characteristics and behavior of visitors of a national park.

3.2. Natural Hazard Analysis

Three of the papers in the Special Issue relate to VGI and natural hazards. The first is a review paper by Klonner et al. [13], who undertook a systematic literature review on research covering VGI and two aspects of disaster management, i.e., preparedness and mitigation. Only 11 papers were included in the analysis since most of the published research focuses on disaster response, highlighting a clear gap where further research is needed, which has also been echoed recently by McCallum et al. [14]. The results showed that most studies were confined to Europe or North America but that a range of natural hazards were dealt with, in particular flooding and forest fires. Unlike with much of VGI, the majority of the studies included the third level of citizen engagement, i.e., participatory, based on Haklay's [15] typology. In contrast, much of VGI and citizen science tends to

be focused on the first level, i.e., crowdsourcing and data collection. Finally, the authors identified a series of challenges including VGI data quality, the need to consider social and not just physical vulnerability, and the current lack of a theoretical framework for integrating different sources of information including VGI, which could be applied more generically to different locations.

This need for integration is the subject of the second paper by Luchetti et al. [16], who provide an example of an early warning system called Whistland. Intended for civil protection and emergency management, the system integrates crowd mapping, Twitter, sensor networks, and Augmented Reality (AR). The architecture of the system is outlined, which consists of a GeoData Collector for accessing Twitter data, a mobile application to visualize the different data sources with AR and a 3D model, and an Analytics Dashboard to query events over time and view heat maps of tweets. The system can be used both in real-time and to visualize different civil protection scenarios for planning purposes. Such an integrated system takes advantage of new citizen and sensor-based data streams for improved decision-making and is likely to be the start of many more developments in this direction.

The final paper by Sosko and Dalyot [17] demonstrates how VGI from mobile phones can be used to augment authoritative weather data, resulting in a denser network and more detailed spatial distributions of weather variables. The focus of the paper is on two variables, i.e., ambient temperature and relative humidity, both of which are important for wildland fire early warning systems. These variables are recorded by the WeatherSignal app using the sensors embedded in smartphones. The paper compares the data collected by volunteers with authoritative reference data and shows that the results are within acceptable margins of error. They also developed algorithms to remove unreliable readings and determined the number of readings needed at a given location to produce stable values. Finally, they interpolated the crowdsourced data and compared this with a weather map created using only the sparser, authoritative network. From this, they demonstrated that the crowdsourced data can provide very useful information on the local patterns, which become visible through the denser, crowdsourced network. Hence, VGI in this context could be a useful source of information for early warning purposes.

3.3. Land Cover Mapping

A further application area in which VGI can serve as a data source is land cover mapping, which is the focus of two papers in this Special Issue. In the first paper, Antoniou et al. [18] investigate the feasibility of geo-tagged photographs for calibrating, validating, and verifying land cover maps. In a review of current applications, the authors first examined the current protocols for geo-tagged photographs and provided an inventory of the metadata collected with regards to Flickr, Panoramio, and Geograph. Photographs were then downloaded from the three sources for the study area of London, and qualitatively evaluated against the metadata requirements for the use cases of land cover calibration, validation, and verification. The authors then manually evaluated the usability of a sub-sample of 3000 photographs. Overall, the results of this study demonstrate the clear potential of all three data sources for all three use cases.

In the second paper by Fonte et al. [19], a procedure is presented that converts OSM data to land cover and compares the results for Kathmandu, Nepal and Dar es Salaam, Tanzania with a recent 30-m global land cover map called GlobeLand30. The two land cover maps are then combined to produce a more up-to-date land cover map for these two cities, which improves the overall accuracy when compared with validation data. The single urban class from GlobeLand30 is then broken down into more detailed classes, using the richness of OSM to populate the map. The paper shows how OSM, in combination with a global land cover map, can be used to create land cover and land use maps in areas where this information may not be readily available.

Table 1. Main themes of the papers in the Special Issue.

Characteristics of VGI			Applications of VGI		
Quality Assessment	Quality Assurance	Review of Existing Approaches	Analysis of Human Activities	Natural Hazard Analysis	Land Cover Mapping
Davidovic et al. [3]	Gómez-Barrón et al. [6]	See et al. [2]	Wang et al. [11]	Klonner et al. [13]	Antoniou et al. [18]
Mahabir et al. [4]	Leibovici et al. [7]		Heikinheimo et al. [12]	Luchetti et al. [16]	Fonte et al. [19]
Touya et al. [5]	Mooney et al. [8]			Sosko and Dalyot [17]	
	Ali et al. [9]				
	Bordogna et al. [10]				

4. The Present and Future of VGI

VGI is the blending of social, economic, and technological factors; it begins with self-motivation and volunteerism and has led to new spatial products and services. Map-making and access to geographic information (GI), which has traditionally been in the realm of governments or industry, has now diffused to a much broader audience. This has led to a step change in the production of GI, which has essentially democratized this process. It is, therefore, not surprising that VGI has become a rapidly growing area of research. Although the present state of VGI cannot be comprehensively covered by a Special Issue, the papers published here give a good overview of the characteristics of this novel data stream, some of the main challenges of using VGI, and a range of possible applications.

The Special Issue continues to show the importance of concerns over the quality of VGI, which is heterogeneous and has different kinds of spatial, temporal, and social bias. The trend towards automated methods, filtering systems, and stronger protocols, as shown in the papers in this Special Issue, may help to alleviate some of these quality concerns in the future. We are not yet in a position where VGI can replace proprietary and authoritative data sets, although OSM continues to be at the forefront of what is possible. We need more examples of such innovative and successful endeavors. Many of the application-oriented papers show the potential of VGI across multiple domains. The importance of VGI for disaster-related applications is a continuing theme that will definitely expand in the future since aiding humanitarian causes is a power incentive for participation. VGI for human behavioral analysis and land cover/land use mapping are only two growing application areas for VGI among many others. The focus on application-oriented research is also clearly going to continue in the future as novel applications of VGI are developed.

Even though VGI was only coined by Goodchild just over 10 years ago, the impact and influence of this phenomenon has been enormous within the spatial domain and beyond. Continuing advances in technology that facilitate participation and the trend towards more inclusive societies where active citizens can play a greater role are omens for exciting times ahead in the field of VGI.

Author Contributions: All authors contributed to the writing of the editorial.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CLC	CORINE Land Cover
CMPA	Column Marginal Proportions of Agreement

COST	Cooperation in Science and Technology
DCDF	Data Conflation or Data Fusion
GL30	GlobeLand30
GNSS	Global Navigation Satellite System
HOT	Humanitarian OpenStreetMap Team
ICT	Information and Communication Technologies
IT	Information Technology
KLL	Kathmandu Living Labs
LBSN	Location Based Social Networks
LULC	Land Use Land Cover
MMU	Minimum Mapping Unit
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
OAc	Overall Accuracy
PAc	Producer's Accuracy
POI	Point of Interest
QA	Quality Assurance
RMPA	Row Marginal Proportions of Agreement
SDGs	Sustainable Development Goals
UA	Urban Atlas
UAc	User's Accuracy
VGI	Volunteered Geographic Information
WPS	Web Processing Service

References

1. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221, doi:10.1007/s10708-007-9111-y.

2. See, L.; Mooney, P.; Foody, G.; Bastin, L.; Comber, A.; Estima, J.; Fritz, S.; Kerle, N.; Jiang, B.; Laakso, M.; et al. Crowdsourcing, Citizen Science or Volunteered Geographic Information? The current state of crowdsourced geographic information. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 55, doi:10.3390/ijgi5050055.

3. Davidovic, N.; Mooney, P.; Stoimenov, L.; Minghini, M. Tagging in Volunteered Geographic Information: An analysis of tagging practices for cities and urban regions in OpenStreetMap. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 232, doi:10.3390/ijgi5120232.

4. Mahabir, R.; Stefanidis, A.; Croitoru, A.; Crooks, A.; Agouris, P. Authoritative and Volunteered Geographical Information in a developing country: A comparative case study of road datasets in Nairobi, Kenya. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 24, doi:10.3390/ijgi6010024.

5. Touya, G.; Antoniou, V.; Olteanu-Raimond, A.-M.; Van Damme, M.-D. Assessing crowdsourced POI quality: Combining methods based on reference data, history, and spatial relations. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 80, doi:10.3390/ijgi6030080.

6. Gómez-Barrón, J.-P.; Manso-Callejo, M.-Á.; Alcarria, R.; Iturrioz, T. Volunteered Geographic Information system design: Project and participation guidelines. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 108, doi:10.3390/ijgi5070108.

7. Leibovici, D.; Rosser, J.; Hodges, C.; Evans, B.; Jackson, M.; Higgins, C. On data quality assurance and conflation entanglement in crowdsourcing for environmental studies. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 78, doi:10.3390/ijgi6030078.

8. Mooney, P.; Minghini, M.; Laakso, M.; Antoniou, V.; Olteanu-Raimond, A.-M.; Skopeliti, A. Towards a protocol for the collection of VGI vector data. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 217, doi:10.3390/ijgi5110217.

9. Ali, A.; Sirilertworakul, N.; Zipf, A.; Mobasheri, A. Guided classification system for conceptual overlapping classes in OpenStreetMap. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 87, doi:10.3390/ijgi5060087.

10. Bordogna, G.; Frigerio, L.; Kliment, T.; Brivio, P.; Hossard, L.; Manfron, G.; Sterlacchini, S. "Contextualized VGI" Creation and management to cope with uncertainty and imprecision. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 234, doi:10.3390/ijgi5120234.

11. Wang, Y.; Jiang, W.; Liu, S.; Ye, X.; Wang, T. Evaluating trade areas using social media data with a calibrated Huff model. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 112, doi:10.3390/ijgi5070112.

12. Heikinheimo, V.; Minin, E.D.; Tenkanen, H.; Hausmann, A.; Erkkonen, J.; Toivonen, T. User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 85, doi:10.3390/ijgi6030085.

13. Klonner, C.; Marx, S.; Usón, T.; Porto de Albuquerque, J.; Höfle, B. Volunteered Geographic Information in natural hazard analysis: A systematic literature review of current approaches with a focus on preparedness and mitigation. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 103, doi:10.3390/ijgi5070103.

14. McCallum, I.; Liu, W.; See, L.; Mechler, R.; Keating, A.; Hochrainer-Stigler, S.; Mochizuki, J.; Fritz, S.; Dugar, S.; Arestegui, M.; et al. Technologies to support community flood disaster risk reduction. *Int. J. Disaster Risk Sci.* **2016**, *7*, 198–204, doi:10.1007/s13753-016-0086-5.

15. Haklay, M. Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing Geographic Knowledge*; Sui, D., Elwood, S., Goodchild, M., Eds.; Springer: Dordrecht, The Netherlands, 2013; pp. 105–122.

16. Luchetti, G.; Mancini, A.; Sturari, M.; Frontoni, E.; Zingaretti, P. Whistland: An augmented reality crowd-mapping system for civil protection and emergency management. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 41, doi:10.3390/ijgi6020041.

17. Sosko, S.; Dalyot, S. Crowdsourcing user-generated mobile sensor weather data for densifying static geosensor networks. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 61, doi:10.3390/ijgi6030061.

18. Antoniou, V.; Fonte, C.; See, L.; Estima, J.; Arsanjani, J.; Lupia, F.; Minghini, M.; Foody, G.; Fritz, S. Investigating the feasibility of geo-tagged photographs as sources of land cover input data. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 64, doi:10.3390/ijgi5050064.

19. Fonte, C.C.; Minghini, M.; Patriarca, J.; Antoniou, V.; See, L.; Skopeliti, A. Generating up-to-date and detailed land use and land cover maps using OpenStreetMap and GlobeLand30. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 125, doi:10.3390/ijgi6040125.