The Relevance of Protocols for VGI Collection

Marco Minghini*, Vyron Antoniou†, Cidália Costa Fonte‡, Jacinto Estima§, Ana-Maria Olteanu-Raimond¶, Linda See‖, Mari Laakso**, Andriani Skopeliti††, Peter Mooney‡‡, Jamal Jokar Arsanjani§§, Flavio Lupia¶¶

*Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy, marco.minghini@polimi.it
†Hellenic Army General Staff, Geographic Directorate, PAPAGOU Camp, Mesogeion 227-231, Cholargos, 15561, Greece
‡Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal / INESC Coimbra, Rua Sílvio Lima, Pólo II, 3030-290 Coimbra, Portugal
§NOVA IMS, Universidade Nova de Lisboa, 1070-312, Lisbon, Portugal
¶Paris-Est, LASTIG COGIT, IGN, ENSG, F-94160 Saint-Mande, France.
‖International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, 2361 Laxenburg, Austria
**Finnish Geospatial Research Institute, Kirkkonummi 02430, Finland
††School of Rural and Surveying Engineering, National Technical University of Athens, 9 H. Polytechniou, Zografou, 15780, Greece
‡‡Department of Computer Science, Maynooth University, Maynooth, Co. Kildare, Ireland
§§Department of Planning and Development, Aalborg University Copenhagen, A.C. Meyers Vænge 15, DK-2450 Copenhagen, Denmark
¶¶Council for Agricultural Research and Economics (CREA), Via Po, 14 00198 Roma, Italy

---

Abstract

Volunteered Geographic Information (VGI) has become a rich and well established source of geospatial data. From the popular OpenStreetMap (OSM) to many citizen science projects and social network platforms, the amount of geographically referenced information that is constantly being generated by citizens is burgeoning. The main issue that continues to hamper the full exploitation of VGI lies in its quality, which is by its nature typically undocumented and can range from very high quality to very poor. A crucial step towards improving VGI quality, which impacts on VGI usability, is the development and adoption of protocols, guidelines and best practices to assist users when collecting VGI. This chapter proposes a generic and flexible protocol for VGI data collection, which can be applied to new as well as to existing projects regardless of the specific type of geospatial information collected. The protocol is meant to balance the contrasting needs of providing VGI contributors with precise and detailed instructions while maintaining and growing the enthusiasm and motivation of contributors. Two real-world applications of the protocol are presented, which guide the collection of VGI in respectively the generation and updating of thematic information in a topographic building database; and the uploading of geotagged photographs for the improvement of land use and land cover maps. Technology is highlighted as a key factor in determining the success of the protocol implementation.

1 Introduction and Background

Volunteered Geographic Information (VGI) represents an important new source of citizen-contributed data (Goodchild, 2007), as outlined in detail in Chapter 2 (See et al., 2017). VGI can be a complementary source of information to authoritative data such as detailed road networks and building footprints, and may be the only source of map data usable after a natural disaster or crisis event has occurred, for example in the case of mapping efforts by the Humanitarian OpenStreetMap Team (HOT)[1]. Yet the main barrier to the widespread use of VGI remains the assessment and documentation of data quality (Johnson and Sieber, 2013; Olteanu-Raimond et al., 2017a). This is particularly true when quality compliance is an essential requirement for VGI exploitation, such as for its exploitation by governments, National Mapping Agencies (NMAs), public bodies (fire fighters, civil protection etc.) and private companies, which

make use of geospatial data to take decisions. From this perspective, an analysis of VGI exploitation by NMAs is made in Chapter 13 (Olteanu-Raimond et al., 2017b), while some guidance on VGI data quality assessment is provided in Chapter 7 (Fonte et al., 2017). The latter chapter describes measures and indicators that are generally applied to VGI after the data have been collected. Instead, more attention should be placed on how to ensure high-quality data collection during the data capture phase. One approach for doing this is to develop and adopt generic and flexible guidelines, best practices and protocols for VGI collection. While guidelines and best practices refer to a set of rules, instructions, suggestions, recommendations or situations that indicate how VGI should be collected, perhaps by reference to examples or ideal cases, protocols can be defined as strict sequences of instructions regulating VGI collection. Specific attention should be paid to the structure and complexity of such guidelines, best practices and protocols; in particular, they should not discourage citizens from contributing, while simultaneously ensuring that the collected data are of an acceptable quality for the purpose of the specific VGI project. Not secondarily, they should ease or facilitate the reuse of VGI for projects and applications other than the one(s) it was originally collected for.

The relevance of establishing protocols in VGI projects and the potential problems for communities and society that arise when these protocols are absent have been highlighted by many authors, including Sui (2007), Johnson and Sieber (2013) and See et al. (2016). In Europe, only a few NMAs have experience with using or integrating VGI in their authoritative datasets (Olteanu-Raimond et al., 2017a), while protocols for VGI within NMAs, governments or Commercial Mapping Companies (CMCs) are lacking (Johnson and Sieber, 2013). Conversely, as mentioned above, many authors have developed methodologies to study the quality of VGI (after it has been collected) and have undertaken VGI comparison, integration or conflation with data from NMAs and CMCs to build more up-to-date, accurate and complete datasets (Girres and Touya, 2010; Haklay, 2010; Ludwig et al., 2011; Al-Bakri and Fairbairn, 2012; Du et al., 2012; Pourabdollah et al., 2013; Touya et al., 2013; Gao et al., 2014; Jokar Arsanjani et al., 2015b; Brovelli et al., 2016a; Fan et al., 2016).

To instruct users in the production of data that are fit-for-purpose, some VGI projects provide detailed guidelines instead of defining a real protocol. OpenStreetMap (OSM)[2] is the most popular VGI project and one of the most studied in the literature (Jokar Arsanjani et al., 2015c); it is extensively described in Chapter 3 (Mooney and Minghini, 2017). Over its more than ten years of life, there has been a progressive development of guidelines about the types of geographic features that users can create and the attributes (or tags) that can be attached to them. The updated version of these guidelines is maintained in a page[3] on the OpenStreetMap Wiki, while their development and enrichment over time is discussed in Chapter 8 (Antoniou and Skopeliti, 2017). It is worth mentioning that, although a real, strict protocol for creating OSM data does not exist and indeed there is considerable freedom left to the contributors,

several studies have documented the high quality of OSM crowdsourced datasets (see e.g. Neis et al., 2011; Fan et al., 2014; Dorn et al., 2015; Jokar Arsanjani et al., 2015a). Another example of VGI project that provides guidelines is the National Map Corps[4], a mapping crowdsourcing programme similar to OSM that supports the Geospatial Information Office of the U.S. Geological Survey (USGS) in gathering rapidly-changing landscape feature data for The National Map (Bearden, 2007).

In other cases, protocols have been designed to assist volunteers in contributing high-quality data that could fit the VGI project's needs and purposes. A well known example is that of Geo-Wiki (Fritz et al., 2012), which is an online crowdsourcing platform where volunteers – provided with a strict and detailed protocol – are asked to use very fine spatial resolution imagery to gather information on land cover and land use to improve global land cover maps. Similarly, an extensive and detailed protocol for digitising old French maps was created and enriched through user collaboration on a dedicated platform[5], which allowed for consistent data records to be maintained (Perret et al., 2015). In the same way, the GéoPeuple project used protocols to create topographic vector datasets from old French maps for analysing population growth (Ruas et al., 2014). The Degree Confluence Project[6] is an example of a project applying a protocol to collect photographs of the landscape from all the intersection points (or confluences) of one degree latitude-longitude around the globe. Volunteers are asked to take either photographs in the four cardinal compass directions (north, south, east, west) or one or more panoramic views from the intersection, one general photograph taken within 100 metres of the confluence, and one photograph of the GPS used. Users then upload all the photographs, along with a text describing the landscape as well as their journey to the confluence point (Fritz et al., 2009). In principle, these photographs may then be reused in another VGI project to yield reference data for map validation (Foody and Boyd, 2012).

The addition of such protocols in VGI projects usually comes with trade-offs; in other words, as the complexity or length of the protocol increases, the participation or retention rate may become lower (see Chapter 5 (Fritz et al., 2017) on motivation and participation for examples). A contrary example to the Degree Confluence Project in the same domain of VGI photograph-based initiatives is represented by Flickr and Panoramio. These are VGI photograph sharing sites that do not provide any protocols regarding how the photographs should be taken or what information should be added. Users can add a title, a comment/description, one or more tags and the location, but these are optional. The lack of protocols is reflected in the very high participation rates (Michel, 2015; Panorank, 2016), but also in the variable quality of the contributions when considering them for applications such as land cover and land use mapping (see e.g. Leung and Newsam, 2012; Estima and Painho, 2014; Antoniou et al., 2016).

To show an example of the variability of the photographs in terms of tags, a random sample of around 130,000 geotagged photographs that were uploaded

to Flickr and Panoramio for the London region in May 2015 was analysed. The frequency of the number of tags associated with the photographs was computed and plotted in Figure 1 as a function of increasing numbers of tags. Clearly the vast majority of photographs (almost 1/3 of the total) have no tags associated to them. In addition, the number of photographs with one to seven tags are within the limits of random variation (although some trends can be spotted; for instance if a user decides to include tags, they usually prefer to append from two to six tags instead of just one). Conversely, the frequency of photographs with eight or more tags shows an almost progressive decrease. This can be seen as a proxy for the following relationship: the more freedom users have in terms of contributions, the more heterogeneous the contributions will be, accompanied with a likely decrease in average quality in terms of their use in further applications. Hence the role of guidelines and protocols could substantially increase the exploitation of VGI for applications not even considered by the person collecting the data.

The definition of protocols is more common in other established citizen science activities where many examples can be found. Accurate data collection by citizens depends on the provision of three elements: clear data collection protocols, simple and logical data forms, and support for participants on protocol use and information submission (Bonney et al., 2009). Pocock et al. (2014) argue that volunteers are more likely to provide information following a given standard if the value of their contribution is recognised. However, if the project requires a complex standard for gathering data, strategies for supporting participants must be deployed and protocols need to be thoroughly tested (Tweddle et al., 2012). Acknowledgement of participants, even simply demonstrating the usefulness of the data, plays a central role in encouraging participation (Pilz et al., 2006).

As discussed in more detail in Chapter 2 (See et al., 2017), VGI can be collected either actively or passively. While in active projects users collect data in a conscious way, passive data collection happens when contributions are gathered without any active engagement (Haklay, 2013). Similarly, Harvey (2013) has made a distinction between truly volunteered versus contributed geographic information (CGI). While the former refers to data that are collected with permission (such as an edit in the OSM database), the latter refers to data collected as part of an automated, open-ended or uncontrollable process (such as the tracking of mobile phones). Information contributed to a passive VGI project typically demands much more processing to result in meaningful information. It is possible to impose a set of protocols in active VGI, but this is usually not possible when using passive VGI or CGI, where the data volumes are often larger than in active sources and hence the data need to be filtered if they are to be used. For example, Bordogna et al. (2015) demonstrated how input data can be filtered based on minimum quality criteria specified by the user, for example to remove geotagged photographs downloaded from repositories such as Flickr and Panoramio.
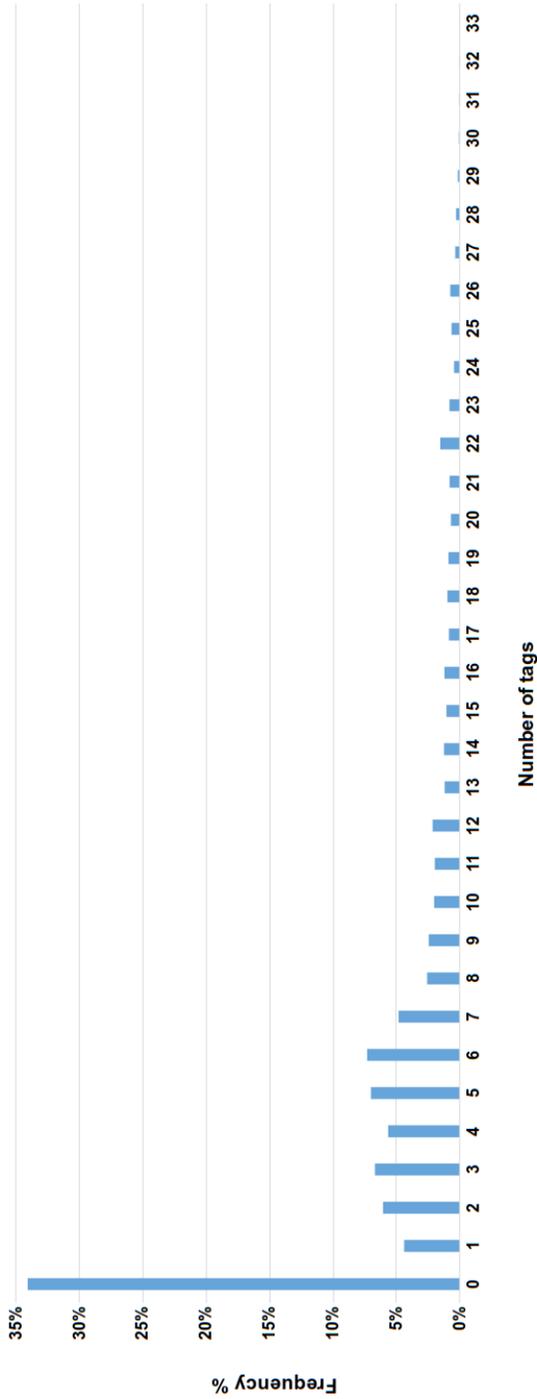
Fig. 1: Frequency of the number of tags for a sample of geotagged photographs uploaded to Flickr and Panoramio (total number of photographs: 130,408; total number of tags: 700,789).

Hence this chapter limits its focus to active VGI projects, where the role played by protocols can be crucial for the quality of the data collected. The chapter seeks to emphasise the need for data collection protocols in VGI projects, and explores how technology can be seamlessly exploited to facilitate collection of suitable data. The chapter takes its origin in a previous work by Mooney et al. (2016), who defined a general and flexible protocol for collecting VGI vector data.

In Section 2 this protocol is briefly presented with the idea of generalising it to all types of VGI projects and VGI data collected. In Section 3 attention is placed on which protocols are required to meet minimum data quality requirements and how technology can play a role in helping to enforce protocols in a user-friendly way. Section 4 presents examples of how the protocol can be applied to two real-world applications, one related to the collection of VGI vector data and the other to geotagged photographs, and reflects upon the relationship between protocols and volunteer motivation. Section 5 concludes the chapter and explores open questions as well as the needs and directions for future research.

2 A Reference Protocol for VGI Collection

A generic protocol has been proposed and developed by Mooney et al. (2016), which can be applied by new VGI projects focused on vector data collection. It can also be used retrospectively on existing data in current VGI projects. This protocol aims to be inclusive of all participants to VGI projects, from new to experienced VGI contributors. By guiding contributors in the process of VGI data collection, the protocol seeks to improve the quality of data in order to both fit the purpose of the specific VGI project for which they are collected and to facilitate their reuse within other, future and potentially unintended, applications. The protocol assumes only a basic working knowledge of geographic information science with basic file and data handling skills from information technology. The protocol has been developed in a bidirectional fashion, i.e. the authors have carefully considered mapping practices in bottom-up approaches (VGI, for example) and top-down approaches (like those used by some NMAs). In this way the protocol is positioned at the intersection between these two opposing approaches for the generation and collection of geographic vector information.

The protocol should be reasonably general and potentially usable by any VGI project based on the collection of vector data through digitisation, field survey or bulk import. The authors have been careful not to relate to any specific VGI initiative, like, for example, OSM, so as to ensure the protocol has potential for further/future customisation or improvement for other specific VGI projects. On the other hand, it gives concrete technical recommendations to easily guide users into a replicable step-by-step data collection process using the tools and

processes that they currently possess and use. The protocol is formalised into five main stages as follows:

- Initialisation
- Data Collection
- Self-Assessment/Quality Control
- Data Submission
- Feedback to the Community.

Initialisation – This involves the users of the protocol becoming familiar with the VGI project and its specific goals and objectives. Familiarisation with the proper devices or technologies for the tasks to be accomplished is required. Users are encouraged to conduct tests of the data collection process to familiarise themselves with the process in general.

Data Collection – Users must carefully plan the data collection process. Data collection in this protocol can be considered as one of the following: digitisation, field survey, or bulk import of existing vector data. Obstacles, problems and technical issues with the specific type of data collection method must be carefully considered before proceeding. At all times data collection must be performed according to the VGI project specifications.

Self-Assessment/Quality Control – This step involves users making their own checks and assessments of their data collection process and the data that have been collected. The users should clearly state if problems were encountered (for instance if there was a GPS signal loss during field collection, licence issues in bulk import, or poor resolution imagery used in digitisation).

Data Submission – In this step users submit, potentially using specific application software, all the data to the project website or application. Submission must be successful and a post-submission check should outline any issues that were encountered during this process.

Feedback to the Community – The protocol encourages users to use all available channels to provide feedback on their experiences. According to Perret et al. (2015), controlling, tracking and reporting all aspects of the process is recommended in VGI. Feedback includes any problems that were encountered, issues that the user resolved, tips or guidance for other users in the project etc.

Despite these five main stages of data collection being intended to be sequential, it is sometimes not easy to establish a well defined limit between them. For example, during data collection the VGI contributors may need to get back to the initialisation stage to get more insight on the project specifications; similarly, contributors may realise that quality control is required again after data submission.

Currently, the protocol described is available to participants in VGI projects in the form of a printed or soft copy manual or document. The future goal of this work is to communicate the concepts of the proposed protocol in order to also influence and guide future software implementations for VGI vector data collection. As will be shown through examples in Section 4, in order for the protocol to be effectively adopted by VGI projects, the role of technology – and hence of VGI software developers – is fundamental. If this protocol can be directly implemented in software within VGI projects, the protocol can be communicated to more users and lead to overall improvements in VGI vector data collection.

3 The Role of Protocols for VGI Quality

While for authoritative data the evaluation of data quality is a well established subject, in VGI it remains rather elusive and vague. What is fundamentally different between authoritative data and VGI is the data collection process. For NMAs and CMCs, rigorous protocols and well defined procedures are in place that must be followed by surveyors. The management of surveyors, the updating of the protocols and the specifications, and the migration from a data scheme to another are fully controlled. A totally different landscape exists for VGI projects, in which the enthusiasm of an enormous but disparate set of volunteers is the driving force. In the case of NMAs and CMCs the logic is simple: production protocols and specifications need to be followed, since the final product will be examined for its quality using various measures (such as the ISO/TC211 quality framework). Similarly, in VGI volunteers should have to fully understand that following or ignoring guidelines, best practices and protocols will have a direct impact on the final spatial product and consequently on its usability. VGI projects can learn a lot from the advances in citizen science. In many cases, the quality of data in citizen science is attained through carefully designed and standardised protocols for participation (Kasperowski and Kullenberg, 2015). Standardisation ensures the validity and accuracy of contributions and classifications performed by citizens (Cohn, 2008: 194). In this context, the following subsections examine, in detail, each of the five data collection stages described above against protocol and best practice instructions.

3.1 Initialisation

One aspect that may influence the quality of the collected information is the type of instructions provided to the volunteers in the *initialisation* stage. While the initial impulse of most trained surveyors is to employ the standard data quality methods from their field, when designing citizen science

projects a different approach for ensuring data quality may be necessary, taking into consideration the degree of participation and the expectations around contributors' skills (Wiggins et al., 2011). If the VGI collection is made for a particular purpose, then the instructions should be detailed enough so that volunteers understand exactly what they are expected to provide. However, instructions with too much detail should be avoided, or at least it should not be mandatory for the volunteer to go through all the detail, because this may be demotivating. The appropriate level of detail of the instructions is, in some circumstances, not easy to establish. Therefore, for some types of VGI projects, studies that identify how volunteers react to several types of instructions should be undertaken, as this reaction may have an important impact on the quality of the generated data (Kerle and Hoffman, 2013). Two practical examples of the importance of instructions for the quality of generated data are the following: if the volunteers need to collect georeferenced photographs, then it should be indicated what must be georeferenced: for example, is it the place where the photograph was taken from or the phenomena shown on the photograph?; and when providing a classification of land cover or disaster damage, how much detailed explanation is required, e.g. the thematic resolution of land cover classes or the choice of one among several damage classes, should be determined.

3.2 Data Collection

Familiarising contributors with the project's aims and goals may enhance their awareness, which, in turn, can help to improve the overall quality of the contributions. Nevertheless, crowdsourced participation inherently suffers from biases, inconsistencies and errors; thus the focus is on how to exclude these inherent characteristics from the ***data collection*** stage. Participation biases can result from various causes. The digital divide, socio-economic factors, demographic distribution and individual perceptions can all have an influence on volunteer contributions (Haklay, 2010; Brovelli et al., 2016b). Here protocols should act preemptively and hinder the appearance of biases. For example, it should be taken for granted that individuals have their own understanding and conceptualisation of the world that might not coincide with a VGI project's mission or specifications. Protocols should clearly state the point of view that volunteers should hold and which processes they should follow to collect the data. In an effort to relieve volunteers from extremely detailed protocols, projects might provide a minimalistic approach on the procedures to follow (Batini et al., 2009). However, this hides two dangers: first, setting the bar lower will probably result in data that are of lower quality. Secondly, more active and experienced volunteers might be discouraged by the approach taken. Thus, the challenge is to provide protocols and best practices that will balance data quality with participation.

3.3 Self-Assessment/Quality Control

Data collection might be influenced by factors that make the process error-prone, leading to errors and inconsistencies in the data. For example, weather, landscape, collaboration with other individuals or the instruments used are just a few factors that might affect in-situ measurements. Here the stage of *self-assessment and quality control* has much to offer. Thus, before uploading data, each volunteer should self-assess the quality of their data and perform all possible quality controls. Protocols should provide enough guidance and explain common pitfalls that can lead to inconsistencies and errors and how to avoid them.

3.4 Data Submission

The next stage for which protocols should provide detailed guidance is *data submission*. Inevitably, individual contributions are generally small, sparse and fragmented, and yet valuable for the evolution of a crowdsourced project. Active and meticulous data collection followed by indifferent data submission (e.g. just pressing the 'upload' button) might not be sufficient. Protocols should stress that data submitted should, when possible, be validated against existing observations or measurements so that no vague or inconsistent cases appear. Even more important is that an individual's work does not harm or destroy other volunteer contributions. This does not mean that updates or alterations should be avoided, but rather that it is important to have a balance between contributor efforts, a way to evaluate the need for change, and a versioning system capable of roll-back to the previous state of the project if needed. Furthermore, submission should not be confined only to data: protocols should require the addition of metadata and supporting/documentation material when possible. For example, filling a form or submitting a geotagged image might be valuable for quality control by other volunteers or moderators. Similarly, any pitfall, problem or simple concern encountered during the data submission stage should be appropriately added to the contributed data.

3.5 Feedback to the Community

Finally, the *feedback to the community* may include the participation in discussion forums, which may help other volunteers to create higher quality data. Perret et al. (2015) highlighted the fact that VGI projects should continuously evolve through the feedback each contributor gets from and gives to others, for instance in terms of how a certain problem encountered while collecting data was solved or any other recommendations or guidance. Communication channels with the VGI project managers and administrators should be provided as

well so that the project itself can evolve based on the user feedback. Thus, a continuous circle is formed that improves the protocol and enhances the overall VGI project quality. This way, common mistakes will hopefully start to disappear and overall data quality will be improved.

4 Applying the Protocol to Real-World Examples

In this section we present two hypothetical, extended examples of real-world applications of the VGI vector data protocol described above. In the first example, the protocol is applied to the updating and collection of new thematic information in a topographic building database. In the second example the protocol is applied to a different domain, that is the collection of photographs for land use / land cover (LULC) mapping.

4.1 Updating and Collecting New Thematic Information in a Topographic Building Database

In this example, an NMA is interested in exploiting crowdsourced vector data to improve their topographic building database. This improvement includes enriching and updating existing building objects (their geometry and thematic information) and capturing new building objects and associated thematic information. Buildings are typically very well mapped by NMAs, but the rapid pace of urban change can mean that keeping their database up-to-date is challenging in terms of resources. Additionally, the thematic information within these databases is often very poor. Typical information which is often missing includes: the function of the building, the number of floors in the building, cultural heritage information related to the building, the entrance(s), etc. As an additional challenge and motivation for VGI contributors, the NMA seeks to create a new layer from scratch to represent the entrances to buildings. This will be a multi-point layer, since a building might have more than one entrance. In this example, the NMA decides to develop a Web-based application to allow citizens to collect data. The implementation and presence of a protocol for this application will greatly assist in reducing the potential submission of low-quality data. Specifically, the Web-based application will use digitisation and field surveys as the means of collecting vector data. The application will present contributors with three layers: a base layer consisting of up-to-date orthoimagery of the region represented in the database; an overlay layer of the existing topographic building object database; and a layer for the entrances to buildings. Contributors will be encouraged to create and/or update the geometry and/or thematic information of building objects to reflect recent changes to building function, structure, etc. Additionally, contributors will be able to add vector point data to building objects to indicate the position of building entrances

along with their door numbers. The implementation of the vector data protocol for this application will ensure that helpful advice and guidance is provided to all contributors in an attempt to maintain and ensure good quality. Guidance is provided for a number of categories:

- *Scale:* Select the appropriate cartographic scale for building level of detail, and preserve it over the collection and contribution process;
- *Shape:* Preserve building shape as much as possible (for instance keep the building corners squared whenever convenient) and digitise minimum details appropriate to the scale;
- *Logical Consistency:* Ensure that new buildings contributed or existing ones that are changed are always closed polygons and do not overlap;
- *Geometric Consistency:* Ensure that multiple entry points to buildings are represented as a multi-point object rather than creating a new point object for each individual entrance in the same building, and that door numbers for each entry point are different;
- *Thematic Quality Control:* Propose a list of thematic attributes and values to the user;
- *Metadata:* Allow free text comments on the visual quality (such as cloud cover, tree cover, shadows or resolution) of the imagery.

The five steps of the protocol workflow outlined in Section 2 are applied to this example as follows:

Initialisation – Citizens will need to register themselves on the Web-based application to use it and contribute vector data and information. Before collecting data, every contributor will need to complete all of the steps in a tutorial demonstration to understand which tasks are required and to familiarise themselves with the processes and tasks in general and with what the goals and objectives of the project are. Depending on the resources available, the NMA may develop a protected 'sandbox' version of the application, where contributors can test out the functionality of the application on a small subset of the topographic buildings database with-out actually making changes to the real database. This form of training will aid learning and help volunteers contribute effectively while still preserving their motivation.

Data Collection – Contributors will be encouraged to carefully plan their collection of new or updated data/information for the application. The application will specifically allow the digitisation of building objects on top of the orthoimagery, the addition of vector point data on building entrances, and the provision of new or updated thematic information associated with building objects. The software application will give prompts and tips to the contributors as they are working.

Self-Assessment/Quality Control – The application will provide functionality to allow contributors to make an initial assessment of the quality of the new data or changes to existing data that they are submitting. For example, if a contributor creates a new building footprint and does not supply any thematic information, the application would indicate this to the contributor. The contributor would then be presented with a generic list of thematic information from which they can choose the appropriate annotations. This would help emphasise the importance of thematic information in the application in the situation where many users may attach greater importance to geometrical data.

Data Submission – In this step, contributors submit their contributed vector data and/or thematic information to the application. The application will provide a space where contributors can provide metadata or descriptive information about their contribution. This could be used by the NMA to assess the overall quality of the contribution, as this information would describe the processes that the contributors used to make their contributions.

Feedback to the Community – The NMA will create a number of information channels to encourage contributors to provide feedback and discussions on their experiences of using the application and contributing vector data using the application. This feedback can include discussions on problems encountered with specific building types or structures, with certain thematic areas, etc. Through these channels, the NMA can provide assistance and feedback to the contributors in the community by offering suggestions on how problems may be fixed or resolved within the application. This creates a complete feedback loop within the vector protocol, which will allow for the protocol to be continuously improved.

4.2   Using Geotagged Photographs for LULC Mapping

In this example, an NMA is interested in exploiting geotagged photographs to improve their LULC maps, and in particular to provide much more data for training their classification algorithms and also to validate the map, if possible. The NMA has already experimented the use of photographs from existing photo-sharing sites such as Flickr and Panoramio, but it was observed that there was too much inconsistency in the tags and in the content of the photographs and thus that not all photographs were usable for the purpose of LULC mapping. Also, there was a strong spatial bias in the distribution of the photographs and not all required LULC types were captured.

Instead, the NMA decides to develop its own national-level photograph-sharing site specifically for the purpose of collecting photographs for LULC mapping, which will have a stricter protocol and ensure higher usable content and tags. At the same time, the data collection protocol should not hamper creativity or the spontaneous enthusiasm that drives contributors while aiming

for the huge volumes of data that are a characteristic of popular social media sites. The NMA decides to develop a customised mobile-based photograph-sharing application, which can use technology to help ensure that specific parts of the data collection protocol are adhered to. The application should have the following features:

- Contributors will be taken through a step-by-step procedure for each location photographed;
- This procedure will require the contributor to take either a set of photographs in the four cardinal directions or a single 360-degree photograph. If the participant chooses the option of taking photographs in four different directions, then the compass in the mobile device will only allow the user to take a photograph when facing the correct cardinal direction;
- The application will prevent participants from using the zoom function, ensuring that the photographs show content closest to their geographic position;
- A 'guide line' will be added to the application so that the contributors can line up the horizon with the 'guide line', so that photographs containing one-third sky and two-thirds landscape are taken;
- The photograph should be dominated by landscape but without restricting the addition of other elements (such as people and animals); moderators or automated methods can be used to assign weights to these photographs for the purpose of LULC creation/validation;
- Once the photographs are taken, the participant will be presented with the possibility to assign tags from a pre-specified list (drawn from the LULC nomenclature used by the NMA) to the photographs, which will be mandatory, along with the possibility to add free form tags, which will be optional;
- The final step in the procedure will be to ask contributors to estimate the distance at which the LULC changes, to indicate how homogeneous or heterogeneous the landscape is;
- There will be at least two modes of operation in the protocol. In the first mode, participants can take photographs at any location, so the geotagged photographs will be useful for creating LULC training datasets; in the second mode, participants will be sent to specific locations, or 'quests' in the form of photograph-caching, which can be used to satisfy the sampling needs of the NMA for LULC map validation and reduce the spatial bias that is common in geotagged photographs from social media photograph-sharing sites.

As much as possible, elements of the protocol will be hidden or incorporated seamlessly into the workflow of the application through technology. In other cases, the protocol will be implemented via elements of gamification, which will be added to maintain, if not grow, the pool of participants and to create a certain level of competition among them, particularly for the photo-caching mode of the application.

Following the vector protocol outlined in Section 2, the five steps are applied as follows:

Initialisation – This first stage will be achieved by providing contributors with a guided tour of the project, including information on how each step contributes to the overall objectives of the project. In addition, step-by-step instructions will be provided to contributors when they first use the application. The guided tour will be mandatory yet short and easy to follow. Once the user has 'passed' through this stage and become familiar with the function of the application, they will be able to take further photographs.

Data Collection – This will be implemented via field survey, which will be facilitated by the mobile application. As outlined above, there will be two main modes of data collection where participants can: (i) photograph landscapes in any location or (ii) be directed to specific locations. Optionally, a third mode will be possible in which participants can turn off the protocol and photograph freely. The purpose of these three modes will be clearly explained to the participants. The mode employed will also allow the NMA to categorise the photographs for a specific use: the first mode may be more suitable for LULC map creation; the second for LULC map validation; while the third can be either omitted or used for training after careful checking.

Self-Assessment/Quality Control – In this step the mobile application will record the positional accuracy and other related parameters (such as dilution of precision (DOP) and type of GPS receiver) as an additional source of information to accompany the photographs. Through the application, the contributor will also estimate the heterogeneity of the LULC, which will provide the NMA with an indication of whether the photograph is in a homogeneous or mixed land cover class. There will be a mechanism implemented that will allow contributors to review the photographs in order to make sure that they comply with the protocol and are of sufficient quality. Contributors will be given the option to retake photographs that are of poorer quality. For instance, in this stage the app will display the position of the photographs taken on top of orthoimagery in order to easily spot positions recorded with low accuracy.

Data Submission – The application will not require data connection in the field but will automatically synchronise the photographs when connected to wifi, so that poor mobile signals will not be an issue. Once photographs are submitted, the online application will allow contributors to view, share and manage their photographs, for instance to correct the tagging of their photographs and thereby improve the labels needed for LULC classification.

Feedback to the Community – The final step will consist in sending out regular information/rich newsletters to contributors, giving them information about levels of improvement in LULC mapping, highlighting those areas that have been better mapped and featuring the contributions of active contributors. It will also highlight what areas are missing and guide

participants to go out and photograph these areas. At this stage, the online application will also allow contributors to rate the contributions of other participants and start conversations and discussions in order to exchange and share suggestions that would lead to an overall improvement in the project's data quality.

Although some research on using geotagged photographs for LULC training and validation has been undertaken in the past (see e.g. Antoniou et al., 2016), this example is still largely hypothetical. However, a similar protocol for collecting geotagged photographs for LULC-related purposes is currently being tested by the FotoQuest Europe student campaign[7]. This initiative asks volunteers to survey specific locations with the purpose of validating the official EU LULC datasets derived from the Land Use and Coverage Area frame Survey (LUCAS) performed by EUROSTAT[8]. For more information on what geotagged photographs can offer, see Chapter 4 (Touya et al., 2017) on using geotagged photographs for examining OSM quality and for verifying the applicability and suitability of various cartographic processes.

5   Discussion and Conclusions

VGI has become a mainstream presence in the GIScience domain. By its own nature, the driving force behind VGI lies in the crowd. The progressive mitigation of the digital divide – not just the traditional one that considers Internet access, but also the second-level digital divide that looks at the real capacity of people to make use of available technology (Hargittai, 2002) – will likely result in an ever increasing amount of contributions uploaded to VGI initiatives. Statistics[9] and predictive models (Jokar Arsanjani et al., 2015a) for the OSM project confirm an increasing growth in both the number of new contributors and submitted data, while Mooney and Winstanley (2015) have argued that VGI contributions can be considered a form of big data. In turn, the increase in VGI may also increase the heterogeneity of contributions and hence solving quality issues for assessing VGI usability may become harder in the future.

In citizen science projects, especially those in the field of conservation and ecology, protocols and guidelines for data collection are generally well developed and clearly accepted by the contributors. In contrast, by its very same nature, the world of VGI has developed in a much freer, diverse and often uncontrolled fashion. Even OSM, which since its birth has dominated the VGI scene, features a culture of freedom in terms of what is mapped and which tags are provided. Hence, this chapter has investigated the need and opportunity to integrate protocols in order to rule and guide the data collection process in active VGI projects, with the purpose of increasing the quality of volunteer contributions. A general and flexible protocol was introduced and described, which can be exploited to standardise data collection processes in

VGI initiatives. The protocol is suitable for implementation in new as well as existing VGI projects and can serve as a reference tool, not just for the project volunteers, but also for the project managers and developers who need to put in place the best possible system to facilitate collection of high-quality data. The implementation of the proposed protocol was illustrated through two different hypothetical examples.

The first example sees an NMA developing an application for crowdsourced data collection aimed at enriching and improving its topographic buildings theme. Data collection includes improving and updating existing building objects (geometry and thematic information) and capturing new features related to buildings and associated thematic information such as entrances. The implementation of the vector data protocol for this application will ensure that helpful advice and guidance is provided to all users in an attempt to maintain and ensure good quality as citizens are contributing changes and new content. The protocol provides guidance on building scale, building shape, logical consistency of building polygon, geometric consistency of entry points to buildings, thematic quality and the provision of metadata. Crucially, the use of a protocol here will allow the NMA to outline guidance on these issues so that high-quality data can be captured. The workflow of the protocol (initialisation, data collection, self-assessment/quality control, data submission and feedback to the community) provides more structure to the contribution process for all users regardless of their background skills or technical abilities.

The second example, an example of implementing the protocol for the collection of geotagged photographs for LULC mapping, involved the hypothetical development of a customised photograph-sharing application by an NMA. However, it could also be beneficial for existing photograph-sharing sites like Flickr and Panoramio to adopt elements of the proposed data collection protocol, recording and providing access to a minimum set of metadata. First, locational information is a common feature of modern mobile phones and some digital cameras, so storing and providing the location as standard information does not present any additional burden to these providers. Moreover, the positional accuracy of handheld devices continues to increase, and there are early efforts to also expand this increased accuracy to indoor positioning (Mautz, 2009; Kuo et al., 2014), so the locational quality of information will continue to become better in the future. Similarly, it could be beneficial to record other elements, such as camera orientation, tilt, etc. These metadata are not only useful for geomatics applications but are also of interest to other domains. A prime example is that of user-contributed tags. From touristic applications (Majid et al., 2013) to early response systems (Masó et al., 2011), tags are considered a semantically rich source of information that need to be further enhanced. Also, the photograph-sharing repositories themselves can gain valuable insights from more complete and rich contributions, since these can be analysed to improve the repositories' own services and attract more participants.

The recognition of the need for protocols to guide future VGI projects is clearly lacking. Hence this chapter has attempted to provide a generic set of guidelines that can help VGI projects consider what elements are necessary to ensure that a minimum data standard is reached while still motivating and sustaining participation. Within this broader project protocol, a protocol for data collection is needed, where we would argue that technology should be used to seamlessly integrate components of the protocol as much as possible, thereby reducing the burden of compliance by contributors. This work provides fruitful ground for future research. The proposed protocol was conceived in a sufficiently general way so that it can be potentially applied to any VGI project. Based on the multiple recommendations and suggestions provided in this chapter, we feel that detailed, customised versions of the protocol can now be created and applied easily to specific VGI initiatives, and that future VGI projects would benefit greatly from adhering to the protocol when designing the data collection process. Applying the protocol to existing or future projects would also serve as a way to determine the value of the protocol itself and to suggest possible improvements. Finally, exploiting the protocol to revise the way in which VGI is collected in a project would allow for the comparison of the quality of data produced before and after the protocol's introduction and therefore to help assess its effectiveness.

Notes

[1] https://hotosm.org
[2] http://www.openstreetmap.org
[3] http://wiki.openstreetmap.org/wiki/Map_Features
[4] http://navigator.er.usgs.gov/help/vgistructures_userguide.html
[5] https://www.geohistoricaldata.org
[6] http://confluence.org
[7] http://www.fotoquest-europe.com
[8] http://ec.europa.eu/eurostat/web/lucas/overview
[9] http://wiki.openstreetmap.org/wiki/Stats

Reference list

Al-Bakri, M., Fairbairn, D., 2012. Assessing similarity matching for possible integration of feature classifications of geospatial data from official and informal sources. *International Journal of Geographical Information Science* 26, 1437–1456. DOI: https://doi.org/10.1080/13658816.2011.636012

Antoniou, V., Fonte, C., See, L., Estima, J., Arsanjani, J., Lupia, F., Minghini, M., Foody, G., Fritz, S., 2016. Investigating the feasibility of geo-tagged

photographs as sources of land cover input data. *ISPRS International Journal of Geo-Information* 5, 64. DOI: https://doi.org/10.3390/ijgi5050064

Antoniou, V and Skopeliti, A. 2017. The Impact of the Contribution Micro-environment on Data Quality: The Case of OSM. In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 165–196. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.h.

Batini, C., Cappiello, C., Francalanci, C., Maurino, A., 2009. Methodologies for data quality assessment and improvement. *ACM Computing Surveys* 41, 1–52. DOI: https://doi.org/10.1145/1541880.1541883

Bearden, M.J., 2007. The National Map Corps. Presented at the Specialist Meeting on Volunteered Geographic Information, University of California at Santa Barbara. Available at: http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Bearden_paper.pdf [Last accessed 1 May 2017].

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., Shirk, J., 2009. Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience* 59, 977–984. DOI: https://doi.org/10.1525/bio.2009.59.11.9

Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M., Rampini, A., 2015. A user-driven selection of VGI based on minimum acceptable quality levels. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-3/W5*, 277–284. DOI: https://doi.org/10.5194/isprsannals-II-3-W5-277-2015

Brovelli, M.A., Minghini, M., Molinari, M., Mooney, P., 2016a. Towards an automated comparison of OpenStreetMap with authoritative road datasets. *Transactions in GIS* 21, 191–206. DOI: https://doi.org/10.1111/tgis.12182

Brovelli, M.A., Minghini, M., Zamboni, G., 2016b. Public participation in GIS via mobile applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 306–315. DOI: https://doi.org/10.1016/j.isprsjprs.2015.04.002

Cohn, J.P., 2008. Citizen science: Can volunteers do real research? *BioScience* 58, 192–197. DOI: https://doi.org/10.1641/B580303

Dorn, H., Törnros, T., Zipf, A., 2015. Quality evaluation of VGI using authoritative data—A comparison with land use data in Southern Germany. *ISPRS International Journal of Geo-Information* 4, 1657–1671. DOI: https://doi.org/10.3390/ijgi4031657

Du, H., Anand, S., Alechina, N., Morley, J., Hart, G., Leibovici, D., Jackson, M., Ware, M., 2012. Geospatial information integration for authoritative and crowd sourced road vector data. *Transactions in GIS* 16, 455–476. DOI: https://doi.org/10.1111/j.1467-9671.2012.01303.x

Estima, J., Painho, M., 2014. Photo based Volunteered Geographic Information initiatives: A comparative study of their suitability for helping quality control of Corine Land Cover. *International Journal of Agricultural and Environmental Information Systems* 5, 73–89. DOI: https://doi.org/10.4018/ijaeis.2014070105

Fan, H., Yang, B., Zipf, A., Rousell, A., 2016. A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. *International Journal of Geographical Information Science* 30, 748–764. DOI: https://doi.org/10.1080/13658816.2015.1100732

Fan, H., Zipf, A., Fu, Q., Neis, P., 2014. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science* 28, 700–719. DOI: https://doi.org/10.1080/13658816.2013.867495

Fonte, C C, Antoniou, V, Bastin, L, Estima, J, Arsanjani, J J, Bayas, J-C L, See, L and Vatseva, R. 2017. Assessing VGI Data Quality. In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 137–163. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.g.

Foody, G., Boyd, D., 2012. Using volunteered data in land cover map validation: Mapping tropical forests across West Africa, in: Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012, pp. 2368–2371. DOI: https://doi.org/10.1109/IGARSS.2012.6352675

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., Obersteiner, M., 2009. Geo-Wiki.Org: The use of crowdsourcing to improve global land cover. *Remote Sensing* 1, 345–354. DOI: https://doi.org/10.3390/rs1030345

Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., Obersteiner, M., 2012. Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software* 31, 110–123. DOI: https://doi.org/10.1016/j.envsoft.2011.11.015

Fritz, S, See, L and Brovelli, M. 2017. Motivating and Sustaining Participation in VGI. In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 93–117. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.e.

Gao, S., Li, L., Li, W., Janowicz, K., Zhang, Y., 2014. Constructing gazetteers from volunteered Big Geo-Data based on Hadoop. *Computers, Environment and Urban Systems*. DOI: https://doi.org/10.1016/j.compenvurbsys.2014.02.004

Girres, J.-F., Touya, G., 2010. Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS* 14, 435–459.

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

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

Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and*

Planning B: Planning and Design 37, 682–703. DOI: https://doi.org/10.1068/b35097

Hargittai, E., 2002. Second-level digital divide: Differences in people's online skills. *First Monday* 7.

Harvey, F., 2013. To volunteer or to contribute locational information? Towards truth in labeling for crowdsourced geographic information, in: Sui, D., Elwood, S., Goodchild, M. (Eds.), *Crowdsourcing Geographic Knowledge*. Springer Netherlands, Dordrecht, Netherlands, pp. 31–42.

Johnson, P., Sieber, R., 2013. Situating the adoption of VGI by government, in: Sui, D., Elwood, S., Goodchild, M. (Eds.), *Crowdsourcing Geographic Knowledge*. Springer Netherlands, pp. 65–81.

Jokar Arsanjani, J., Mooney, P., Helbich, M., Zipf, A., 2015a. An exploration of future patterns of the contributions to OpenStreetMap and development of a Contribution Index: Future Patterns of the Contributions to Open-StreetMap. *Transactions in GIS* 19, 896–914. DOI: https://doi.org/10.1111/tgis.12139

Jokar Arsanjani, J., Mooney, P., Zipf, A., Schauss, A., 2015b. Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets, in: Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (Eds.), *OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography*. Springer International Publishing, Cham, pp. 37–58.

Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (Eds.), 2015c. *Open-StreetMap in GIScience, Lecture Notes in Geoinformation and Cartography*. Springer International Publishing, Cham.

Kasperowski, D., Kullenberg, C., 2015. Learning in citizen science projects: types, behaviour and protocols, in: *The Conference on Computer Science Collaborative Learning, CSCLROWD2015: Designing Futures for Learning in the Crowd: New Challenges and Opportunities for CSCL*, Gothenburg, Sweden, 7 June 2015. Available at: https://csclcrowd2015.files.wordpress.com/2015/04/csclcrowd2015_submission_4.pdf [Last access 16 May 2017]

Kerle, N., Hoffman, R.R., 2013. Collaborative damage mapping for emergency response: the role of Cognitive Systems Engineering. *Natural Hazards and Earth System Science* 13, 97–113. DOI: https://doi.org/10.5194/nhess-13-97-2013

Kuo, Y.-S., Pannuto, P., Dutta, P., 2014. Demo: Luxapose: Indoor positioning with mobile phones and visible light, in: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, Mobi-Com '14. ACM, New York, NY, USA, pp. 299–302. DOI: https://doi.org/10.1145/2639108.2641747

Leung, D., Newsam, S., 2012. Exploring geotagged images for land-use classification, in: Proceedings of the ACM Multimedia 2012 Workshop on Geotagging and Its Applications in Multimedia (GeoMM '12). Presented

at the GeoMM '12, ACM Press, New York, NY, pp. 3–8. DOI: https://doi.org/10.1145/2390790.2390794

Ludwig, I., Voss, A., Krause-Traudes, M., 2011. A comparison of the street networks of Navteq and OSM in Germany, in: Geertman, S., Reinhardt, W., Toppen, F. (Eds.), *Advancing Geoinformation Science for a Changing World*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 65–84.

Majid, A., Chen, L., Chen, G., Mirza, H.T., Hussain, I., Woodward, J., 2013. A context-aware personalized travel recommendation system based on geotagged social media data mining. *International Journal of Geographical Information Science* 27, 662–684. DOI: https://doi.org/10.1080/13658816.2012.696649

Masó, J., Diaz, P., Pons, X., 2011. Mapping geo-tagged pictures in a disaster management event, in: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Antalya, Turkey, 3–8 May 2011, available at: http://www.isprs.org/proceedings/2011/Gi4DM/PDF/OP24.pdf [Last accessed 16 May 2017].

Mautz, R., 2009. Overview of current indoor positioning systems. *Geodezija ir Kartografija* 35, 18–22. DOI: https://doi.org/10.3846/1392-1541.2009.35.18-22

Michel, F., 2015. How many photos are uploaded to Flickr every day, month, year? Available at https://www.flickr.com/photos/franckmichel/6855169886 [Last accessed 27 June 2016].

Mooney, P and Minghini, M. 2017. A Review of OpenStreetMap Data. In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 37–59. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.c.

Mooney, P., Minghini, M., Laakso, M., Antoniou, V., Olteanu-Raimond, A.-M., Skopeliti, A., 2016. Towards a protocol for the collection of VGI vector data. *ISPRS International Journal of Geo-Information* 5(11), 217. DOI: https://doi.org/10.3390/ijgi5110217

Mooney, P., Winstanley, A., 2015. Is VGI Big Data? in: *Proceedings of GISRUK 2015*, extended abstracts, University of Leeds, Leeds, UK, 15–17 April 2015. Available at: http://leeds.gisruk.org/abstracts/GISRUK2015_submission_65.pdf [Last accessed 27 June 2016].

Neis, P., Zielstra, D., Zipf, A., 2011. The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet* 4, 1–21. DOI: https://doi.org/10.3390/fi4010001

Olteanu-Raimond, A.-M., Hart, G., Foody, G., Touya, G., Kellenberger, T., Demetriou, D., 2017a. The scale of VGI in map production: A perspective of European National Mapping Agencies. *Transactions in GIS* 21, 74–90. DOI: https://doi.org/10.1111/tgis.12189

Olteanu-Raimond, A-M, Laakso, M, Antoniou, V, Fonte, C C, Fonseca, A, Grus, M, Harding, J, Kellenberger, T, Minghini, M, Skopeliti, A. 2017b. VGI in National Mapping Agencies: Experiences and Recommendations.

In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 299–326. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.m.

Panorank, 2016. Global stats. Available at http://www.panorank.com/index.php?lang=en&op=global. [Last accessed 27 June 2016].

Perret, J., Gribaudi, M., Barthelemy, M., 2015. Roads and cities of 18th century France. *Scientific Data* 2, 150048. DOI: https://doi.org/10.1038/sdata.2015.48

Pilz, D., Ballard, H.L., Jones, E.T., 2006. Broadening Participation in Biological Monitoring: Handbook for Scientist*s and Managers (General Technical Report No. PNW-GTR-680)*. United States Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA.

Pocock, M.J.O., Chapman, D.S., Sheppard, L.J., Roy, H.E., 2014. *Choosing and Using Citizen Science: a guide to when and how to use citizen science to monitor biodiversity and the environment*. Centre for Ecology & Hydrology, UK.

Pourabdollah, A., Morley, J., Feldman, S., Jackson, M., 2013. Towards an authoritative OpenStreetMap: Conflating OSM and OS openData national maps' road network. *ISPRS International Journal of Geo-Information* 2, 704–728. DOI: https://doi.org/10.3390/ijgi2030704

Ruas, A., Plumejeaud, C., Nahassia, L., Grosso, E., Olteanu, A.-M., Costes, B., Vouloir, M.-C., Motte, C., 2014. GéoPeuple: The creation and the analysis of topographic and demographic data over 200 Years, in: Buchroithner, M., Prechtel, N., Burghardt, D. (Eds.), *Cartography from Pole to Pole*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–17.

See, L, Estima, J, Pőd.r, A, Arsanjani, J J, Bayas, J-C L and Vatseva, R. 2017. Sources of VGI for Mapping. In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 13–35. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.b

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M., Liu, H.-Y., Milčinski, G., Nikšič, M., Painho, M., Pődör, A., Olteanu-Raimond, A.-M., Rutzinger, M., 2016. Crowdsourcing, citizen science or Volunteered Geographic Information? The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information* 5, 55. DOI: https://doi.org/10.3390/ijgi5050055

Sui, D.Z., 2007. Volunteered Geographic Information: A tetradic analysis using McLuhan's law of the media. Position Paper for the NCGIA and Vespucci Workshop on Volunteered Geographic Information. Santa Barbara, CA, 13–14 Dec 2007. Available at: http://ncgia.ucsb.edu/projects/vgi/docs/position/Sui_paper.pdf [Last accessed 27 June 2016].

Touya, G, Antoniou, V, Christophe, S and Skopeliti, A. 2017. Production of Topographic Maps with VGI: Quality Management and Automation. In: Foody, G, See, L, Fritz, S, Mooney, P, Olteanu-Raimond, A-M, Fonte, C C

and Antoniou, V. (eds.) *Mapping and the Citizen Sensor*. Pp. 61–91. London: Ubiquity Press. DOI: https://doi.org/10.5334/bbf.d.

Touya, G., Coupé, A., Jollec, J., Dorie, O., Fuchs, F., 2013. Conflation optimized by least squares to maintain geographic shapes. *ISPRS International Journal of Geo-Information* 2, 621–644. DOI: https://doi.org/10.3390/ijgi2030621

Tweddle, J.C., Robinson, L.D., Pocock, M.J.O., Roy, H., 2012. *Guide to citizen science: developing, implementing and evaluating citizen science to study biodiversity and the environment in the UK*. Natural History Museum and NERC Centre for Ecology & Hydrology.

Wiggins, A., Newman, G., Stevenson, R.D., Crowston, K., 2011. Mechanisms for data quality and validation in citizen science. Presented at the Seventh International Conference on e-Science Workshops, IEEE, pp. 14–19. DOI: https://doi.org/10.1109/eScienceW.2011.27