Huerta, Schade, Granell (Eds): Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Castellón, June, 3-6, 2014. ISBN: 978-90-816960-4-3

# Characteristics of Citizen-contributed Geographic Information

Spyridon Spyratos
University of Thessaly,
Pedion Areos, 38334, Volos, Greece
&
Joint Research Centre, European Commission,
via Enrico Fermi 2749, 21027, Ispra, Italy
spyridon.spyratos@jrc.ec.europa.eu

Michael Lutz
Joint Research Centre,
European Commission,
via Enrico Fermi 2749,
21027, Ispra, Italy
michael.lutz@jrc.ec.europa.eu

Francesco Pantisano
Joint Research Centre,
European Commission,
via Enrico Fermi 2749,
21027, Ispra, Italy
francesco.pantisano@jrc.ec.europa.eu

**Abstract**

Current Internet applications have been increasingly incorporating citizen-contributed geographic information (CCGI) with much heterogeneous characteristics. Nevertheless, despite their differences, several terms are often being used interchangeably to define CCGI types, in the existing literature. As a result, the notion of CCGI has to be carefully specified, in order to avoid vagueness, and to facilitate the choice of a suitable CCGI dataset to be used for a given application. To address the terminological ambiguity in the description of CCGI types, we propose a typology of GI and a theoretical framework for the evaluation of GI in terms of data quality, number and type of contributors and cost of data collection per observation. We distinguish between CCGI explicitly collected for scientific or socially-oriented purposes. We review 27 of the main Internet-based CCGI platforms and we analyse their characteristics in terms of purpose of the data collection, use of quality assurance and quality control (QA/QC) mechanisms, thematic category, and geographic extents of the collected data. Based on the proposed typology and the analysis of the platforms, we conclude that CCGI differs in terms of data quality, number of contributors, data collection cost and the application of QA/QC mechanisms, depending on the purpose of the data collection.

*Keywords*: Volunteered Geographic Information (VGI), Citizen Science, Crowd sourced geographic information, Citizen-Contributed Geographic Information (CCGI), Social Geographic Data (SGD)

## 1 Introduction

Recent social and technological developments, such as the increased educational attainment and the diffusion of sensor-enabled devices increase the number of citizens who are potentially able to collect and publicly share almost real time geographic information (GI) on the Internet. Such a citizen-contributed geographic information (CCGI) differs from GI collected by professionals in the context of professional routines and practices for four main reasons. First, the CCGI data collectors possess significantly diverse level of scientific and technical knowledge [2]. Second, the CCGI data collection methods and equipment are very different and often unknown. Third, the quality of CCGI is not always ensured and controlled by formal quality assurance procedures [14], and, finally, CCGI is mostly collected at time and locations that are generally not defined a priori by an organization.

Lately, an increasing number of Internet-based platforms has been developed with the purpose of collecting CCGI for both socially-oriented and scientific purposes. These platforms consist of hardware and software components, such as servers and mobile application interfaces, as well as analytical tools for data processing. They cover data about various environmental domains, such as acoustic pollution [30], biodiversity [16] and land cover observations [8]. Clearly, since CCGI data is gratuitously contributed by the citizens, these platforms offer timely GI and at very limited cost [11].

Due to these reasons, CCGI is increasingly used as auxiliary input for environmental monitoring and mapping [20, 29] and research studies [7]. However, due to the numerous types of existing CCGI, it is still unclear whether and what types of CCGI can contribute towards a better and more holistic understanding of the environment. Goodchild and Li [11] suggest that volunteered geographic information (VGI) is often inadequate data source for scientific research, because "its quality is highly variable and undocumented, it fails to follow scientific principles of sampling design, and its coverage is incomplete". In contrast, Lee [18] mentions that much of the knowledge about the USA climate is based on long-term volunteer records. In this respect, we argue that both of the above statements are valid, as they refer to different types of CCGI.

In fact, CCGI is not a homogenous category and includes GI that significantly differs in terms of purpose of data collection, data quality and the characteristics of contributors. Nevertheless, in the literature, terms such as VGI [10], crowd sourced geographic information, and user generated geographic content (UGGC) are often being used interchangeably to describe various GI types. For example, VGI describes a distinct subset of CCGI, UGGC and crowd-sourced GI as it embodies the notion of volunteering for data collection [5]. VGI describes a science-oriented phenomenon that is supported by technology. Devising CCGI categories is a fundamental operation, as the definition of each of these categories has to denote the characteristics of the collected data, and the characteristics of the contributors e.g. volunteers or users of social networking applications.

In this study, we address this terminological ambiguity in the description of CCGI types, and we provide guidelines for GI type definition. First, based on the purpose of the data collection activity, we propose a typology of CCGI and we identify factors that affect the data quality and quantity of the collected data. Second, we identify Internet-based platforms that collect CCGI, we classify them based on the proposed typology, and we analyse three characteristics of CCGI platforms and datasets. These characteristics are: (a) the existence of quality control and quality assurance (QA/QC) mechanisms that depend on citizens, (b) the thematic category, and (c) the geographic extent of the collected data.

The main rationale of this work is to propose a theoretical framework for the evaluation of CCGI data to be used for scientific or social applications.

The remainder of the paper is structured as follows. Section 2 describes the proposed typology of CCGI. Section 3 presents the methodology followed for identifying and analysing CCGI platforms and datasets and the results of their analysis. In Section 4, we discuss the results of the analysis. Finally, future work and conclusions are outlined in Section 5.

## 2 Typology of citizen-contributed geographic information

The existing literature includes two CCGI typologies [1, 3] which relevant to the purpose of the current study. The first, proposed by Antoniou et al. [1], introduces a distinction between spatially implicit and explicit UGGC web applications, based on their declared objectives. The second, by Craglia et al. [3], defines four VGI types based on two dimensions which can be either explicit or implicit. These dimensions are "first, the way the information was made available, and second, the way geographic information forms part of it" [3].

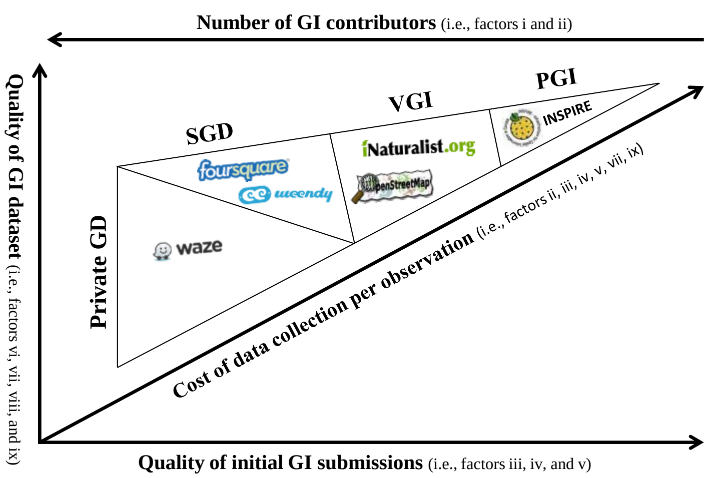
To address the terminological ambiguity in the description of CCGI types, and to support the analysis of platforms, provided in Section 3, we propose a typology of GI which, in contrast to the existing ones, is based on the purpose of the data collection. In the proposed typology (see Fig. 1) we distinguish between CCGI collected for scientific (VGI) and socially-oriented (Social Geographic Data) purposes which are defined as:

- Volunteered Geographic Information (VGI). In this study VGI refers to GI intentionally collected by citizens, in the context of real life or on-line science-oriented voluntary activities. For instance, the VGI category includes GI collected by volunteers as part of a broad scientific enquiry in the data collection stage of citizen science projects (for more details on citizen science see Silvertown [25]) or in the context of crowdsourcing projects [15] e.g. Google Map Maker [12].
- Social Geographic Data (SGD). The SGD category describes geographic or geo-referenced data that is publicly available over the Internet and it has been generated by citizens for socially oriented purposes. For example, this category includes Foursquare place data [6], and geo-located public tweets [28].

Apart from the above CCGI types, two other categories of GI exist:

- Professional Geographic Information (PGI) [22]. PGI is composed by GI exclusively collected by experts, e.g. surveyors or urban planners, in the context of professional routines and practices.
- Private Geographic Data (Private GD) category includes geographic or geo-tagged data that has not been publicly shared by the data author. Private GD is produced by citizens and it can either be data that is associated with the characteristics of an individual or data intended for a particular person, group or service. For example, this category includes not-publicly shared geo-located tweets [28], and Global Navigation Satellite System (GNSS) data contributed to navigation services.

Fig. 1: Typology of GI



This paper focusses on CCGI, i.e., GI collected and publicly shared by citizens. PGI and Private GD are out of the scope of this study, since the former includes only qualified professional in its collection, and the latter deals with data not publicly contributed and not intended to be reused, other than by the initial recipients.

## 2.1 Characteristics of GI datasets

In the proposed typology, we distinguish between three main characteristics (see Fig. 1) for SGD, VGI, PGI and Private GD. The characteristics of the data collection activity, of the GI contributors, platforms and data collection tools, are factors that impact the characteristics of the collected data. These characteristics are: *the number of potential GI contributors*, *the quality of initial GI submission*, *the overall quality of the GI datasets*, and *the cost of data collection*. Due to the scope of this study the analysis is focused on the CCGI, namely the SGD and the VGI.

### 2.1.1. Number of potential GI contributors

As shown in the upper axes of Fig. 1, the number and the demographic profile of citizens that can potentially collect GI depends on the following factors:

i. The level of technical and scientific knowledge required for data collection.
ii. The time, technical equipment and other resources needed for data collection [13].

These two factors limit the number of citizens who can autonomously participate in science-oriented or socially-oriented data collection activities. Regarding the scientific and technical knowledge of VGI data collectors (i.e. factor i), a study by Budhathoki et al. [2] revealed that 25% of the OpenStreetMap contributors had more than 1 year experience with GISystems and the 49% had none. Statistics like these highlight the fact that the demographic profile of VGI data collectors is heterogeneous and not representative of the society. Additionally, such statistics prove that VGI data collectors are not largely untrained, and confirm Lee's [18] statement that volunteer does not necessarily equal amateur.

In contrast to VGI, SGD is not the product of science-oriented tasks, and thus, the level of scientific knowledge required for the collection of SGD observations is, in principle, lower compared to VGI. Thus, SGD can additionally be collected by citizens with low-level science skills. As a result, the number of potential SGD contributors is typically larger compared to the number of VGI contributors.

### 2.1.2. Quality of initial GI submissions

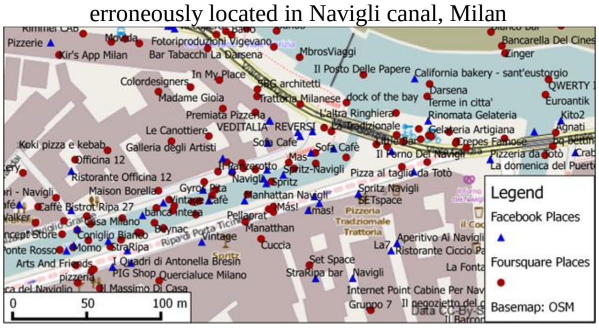
The quality of initial GI submissions refers to the quality of the first GI data submission by a citizen, before any correction or filtering is made by the QA/QC mechanisms. For an extensive survey on the quality elements of GI, such as the positional and thematic accuracy, we refer the interested reader to Oort [21]. As shown in the bottom axes of Fig. 1, the quality of initial GI submissions depends on factors such as:

iii. The desired (or de-facto, de jure) accuracy of GI.
iv. The scientific and technical knowledge of data collectors [4, 24].
v. The accuracy of the utilized equipment, sensors, and auxiliary data, e.g. satellite images.

Factor (iv) relies on the contributors characteristics, while factors (iii), and (v) also depend on the platforms. For instance, for mapping applications, the accuracy of an observation depends both on the accuracy of the GNSS sensors that citizens deploy, and on the quality of the auxiliary satellite images that a platforms provides.

According to our definition, VGI is collected for scientific purposes, and thus, the desired positional and thematic accuracy (i.e. factor iii) and the quality of utilized sensor (i.e. factor v) are both higher compared to SGD. The reason is that a volunteer aims at describing a phenomenon or a feature as accurately as possible. Instead, users of socially-oriented web applications demand a level of accuracy that is sufficient to efficiently convey a geo-tagged message. For example, Fig. 2 shows the Navigli area in Milano, Italy, where many of the Facebook and Foursquare places are mistakenly pinned in the water. The place data positional precision is clearly not suitable for mapping or routing purposes.

Fig. 2: Many Facebook and Foursquare place data are erroneously located in Navigli canal, Milan




Sources: Place data, Facebook Graph API and Foursquare Venues API; Basemap, OSM contributors.

### 2.1.3. Quality of GI datasets

The quality of VGI and SGD significantly varies across time and space, even within the same dataset. As a matter of fact, VGI and SGD datasets are highly heterogeneous, as they are composed by observations that differ in terms of equipment accuracy and citizen technical and scientific background, even in local spatial scale. We note that the overall quality of the GI datasets in a given area mainly depends on the following factors:

vi. The quality of the initial GI submissions.
vii. The number and the demographic profile of contributors and the number of contributions.
viii. The existence and the application of QA/QC mechanisms.
ix. The degree of coordination for the data collection activity.

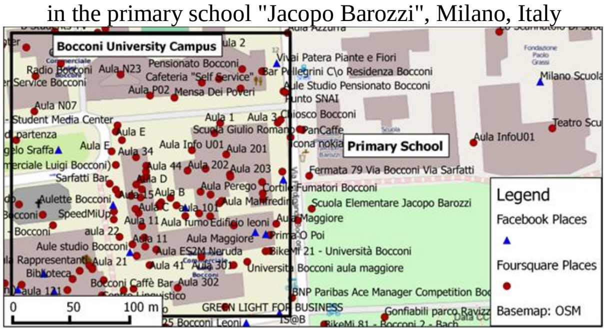
The quality of GI datasets is determined to a great extent by the quality of initial GI submissions (i.e. factor vi) from which are derived. The demographic profile, the number and the spatial distribution of CCGI contributors are factors (factor vii and in more detail see Section 2.1.1) that affect the thematic and spatial completeness of a CCGI dataset [13, 27]. The existence of horizontal or hierarchical coordination of a data collection activity (i.e. factor ix) clearly has a positive impact on the spatial and temporal completeness of a dataset.

QA/QC mechanisms are adopted for the purpose of improving the quality of GI. QA/QC mechanisms can be managed by professionals in the context of professional routines and practices, and/or by the community of contributors, in case citizens assess the correctness of the observations. In addition, QA/QC mechanisms can be supported by automated procedures, in which each observation is automatically checked based on predefined rules, as in [19], for example. In citizen-based QA/QC mechanisms, the quality of the observations stored in the GI datasets depends on the number of contributors (i.e. factor vii), which are also reviewers [9, 14]. This relation directly confirms “Linus’ law” [23], stating that the higher the number of users or contributors of a product is, the higher is the probability that a problem will be fixed by someone.

Several studies have proved that the overall quality of VGI datasets is inferior to PGI [9, 13, 17]. However, few studies have addressed the quality of SGD. The Antoniou e. al. [1] study demonstrate that the spatial distribution of SGD observations is more likely to be limited to the users’ existing activity space compared to VGI spatial distribution. SGD is collected in the context of the data collectors’ social activities, and not as part of a scientific inquiry. For this reason, VGI datasets are expected to have higher spatial and temporal completeness, compared to SGD.

For instance, Fig. 3 shows Foursquare and Facebook place data in an area of Milan, Italy. On the left side of Fig.3, the Bocconi University is well covered while a primary school on the right side is not. The reason for this is that only a limited number of primary school students or staff are declaring the physical presence on Facebook or Foursquare. As a result, their activity space is not well covered on Facebook and Foursquare place datasets.

Fig. 3: Abundance of Facebook and Foursquare place data in a detailed level in Bocconi University campus at the left side of the figure, versus scarcity of place data at the right side, e.g., in the primary school "Jacopo Barozzi", Milano, Italy




Sources: Place data, Facebook Graph API and Foursquare Venues API; Basemap, OSM contributors.

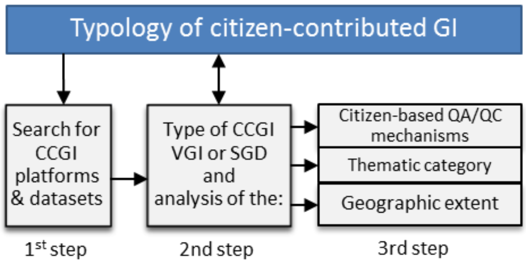
### 2.1.4. Cost of data collection per observation

The financial cost of data collection and processing per observation is another important characteristic of GI. Factors that affect this financial cost are ii, iii, iv, v, viii, and ix. In principle, the higher the quality of the technical and human resources used for data collection, the higher the cost for their usage is. For example, professional GNSS receivers are more accurate and expensive than those built-in mobile phones [31]. The application of QA/QC mechanisms, and the efforts made for coordination of the data collection activity are also factors that have a considerable financial cost for data collection. As a matter of fact, each GI type incurs different costs for data collection. For the collection of PGI, a professional staff is hired, while for the VGI and SGD the contributors are volunteers. Professional trainers are commonly used to train PGI and VGI data collectors, while this is not the case for SGD and Private GD. It is, therefore, arguable that SGD is less expensive to collect than VGI and PGI.

## 3 Methodology & Results

In this section, we focus our analysis on Internet-based platforms that collect CCGI about environmental elements, such as atmosphere, water, soil, land and landscape. We decided to analyse CCGI platforms, in an effort to study how the purpose of the data collection affects the characteristics of the collected CCGI datasets. The methodology for identifying and analysing CCGI platforms and datasets is presented in Fig. 4.

Fig. 4: Methodology followed for identifying and analysing CCGI platforms and datasets




The first step of the methodology was the identification of CCGI platforms that collect data on the environmental elements. For the identification of these platforms an extensive search of the English literature and Web resources was conducted. The searches were performed by using English keywords, which are typically used to describe CCGI. These terms and their variants are:

a) Volunteered geographic/environmental information/data
b) User-generated geographic/spatial content.
c) Crowd sourced geographic/environmental information/ data.

During the search period, 27 platforms (see Table 1) were identified. Given the method for identifying the platforms, the results mostly include popular English-based platforms. Therefore, the results of the platforms analysis cannot be quantitatively generalized, but could be used for understanding the CCGI characteristics.

The second step of the methodology was the analysis of the type of CCGI that the 27 platforms collect. Based on the proposed typology, we classified the 27 platforms into VGI and SGD (see third column of Table 1). The reason for this is that the purpose of the data collection, as defined by each platform's objectives, affects the characteristics of the collected data, such as, its spatial distribution and its accuracy.

Table 1: Name and type and website of CCGI platforms

| No | Name of platform | Type of GI | Website |
|---|---|---|---|
| 1 | Aircasting | VGI | aircasting.org |
| 2 | AirProbe | VGI | cs.everyaware.eu/event/airprobe |
| 3 | ARGO Sentinel | VGI | argomobile.isti.cnr.it |
| 4 | CWOP | VGI | wxqa.com |
| 5 | Facebook Places | SGD | www.facebook.com |
| 6 | FishBase | VGI | www.fishbase.org |
| 7 | Flickr | SGD | www.flickr.com |
| 8 | Foursquare Venues | SGD | foursquare.com |
| 9 | Geograph | VGI | www.geograph.org.uk |
| 10 | Geowiki | VGI | www.geo-wiki.org |
| 11 | Google Map Maker | VGI | www.google.com/mapmaker |
| 12 | iNaturalist | VGI | www.inaturalist.org |
| 13 | iRecord | VGI | www.brc.ac.uk/irecord |
| 14 | iSPEX | VGI | ispex.nl/en |
| 15 | iSpot | VGI | www.ispot.org.uk |
| 16 | NoiseTube | VGI | www.noisetube.net |
| 17 | Noisewatch | VGI | eyeonearth.org/map/NoiseWatch |
| 18 | OpenStreetMap | VGI | openstreetmap.org |
| 19 | Panoramio | SGD | www.panoramio.com |
| 20 | PSW Weather | VGI | www.pwsweather.com |
| 21 | The National Map Corps | VGI | navigator.er.usgs.gov |
| 22 | WaterWatch | VGI | eyeonearth.org/map/waterwatch |
| 23 | Weathersignal | VGI | weathersignal.com |
| 24 | WeatherUnderground | VGI | www.wunderground.com |
| 25 | Weendy | SGD | www.weendy.com |
| 26 | Wheel Map | VGI | www.wheelmap.org |
| 27 | WideNoise | VGI | cs.everyaware.eu/event/widenoise |

The third step of the methodology included the analysis of three characteristics of CCGI platforms and datasets. The first characteristic that we analysed is the type of QA/QC mechanisms which depend on citizens. Citizen-based QA/QC mechanisms allow the users of the platforms to review and rate the correctness of VGI and SGD observations. Citizen-based QA/QC mechanisms can vary from being horizontally structured, in which user have distributed and equal authorities on editing observations, to more hierarchically structured, in which community representatives or elite users have increased editing authorities compared to average users.

There are two types of citizen-based QA/QC mechanisms. The first type allows citizens to edit an observation or suggest an edit to its author. The second type allow citizens to rate the accuracy of an observation, and thus, to also assess the competence of the data contributor. Based on the existence and the type of citizen-based QA/QC mechanisms, we classified the 27 platforms in four categories (see Table 2).

Table 2: Citizen-based QA/QC mechanism of CCGI platforms

| Citizen-based QA/QC | VGI Platforms | SGD platforms |
|---|---|---|
| None | 1; 2; 3; 4; 13; 14; 16; 17; 20; 22; 23; 24; 27 | 7; 19; 25 |
| Only rate | None | None |
| Only edit | 6; 9; 10; 11; 18; 21; 26 | 5; 8 |
| Rate and edit | 12; 15 | None |

The second characteristic that we analysed was the thematic category of the data that the 27 platforms collect. We used a context based classification into six thematic categories as shown in Table 3. Moreover, we classified the six thematic categories into two groups. The first includes CCGI about continuous geographic phenomena and the second CCGI about discrete geographic features.

Table 3: Thematic category of CCGI datasets

| | Thematic category | VGI platforms | SGD platforms |
|---|---|---|---|
| Phenomena | Noise | 1; 16; 17; 27 | None |
| | Meteorology | 1; 2; 4; 20; 23; 25 | 25 |
| | Air quality | 1; 2; 14 | None |
| | Water quality | 3; 22 | None |
| Features | Biodiversity, species occurrences | 6; 12; 13; 15; | 7; 19 |
| | Topography, place, land cover and landscape | 9; 10; 11; 18; 21; 26 | 5; 7; 8; 19 |

Finally, we analysed the geographic extent of CCGI datasets. The geographic extent can be local, national, multi-national, or global. As shown in the Table 4, the geographic extent of the most CCGI data sources that were identified in this study is global.

Table 4: Geographic extent of CCGI datasets

| Geographic extent | VGI Platforms | SGD platforms |
|---|---|---|
| Global | 1; 2; 3; 4; 6; 10; 11; 12; 15; 16; 17; 18; 20; 23; 24; 26; 27 | 5; 7; 8; 19; 25 |
| Multi-National | 9 (UK, IL); 22 (EU) | None |
| National | 13(UK); 21(US); 14(NL) | None |
| Local | None | None |

## 4 Discussion

SGD and VGI are collected in the context of socially and science oriented activities respectively. As we have discussed in Section 2, SGD and VGI differ in terms of the quality of initial GI submissions, the overall quality of GI datasets, the number of the potential contributors, and the data collection cost per observation. Although SGD is collected for socially-oriented purposes, it can be reused in the context of scientific applications. An example is given by the reuse of Panoramio photos as auxiliary input for land cover mapping [27].

Most of the VGI platforms and all of the SGD platforms analysed in this study have global geographic extent. The reason is that the identified platforms are biased towards popular, and due to their popularity are more likely to be used by users and volunteers worldwide. The development and maintenance of CCGI platforms is a task that requires significant financial resources and technical skills. Thus, local participatory data collection and citizen science initiatives are more likely to use existing well-established CCGI Internet platforms for collecting data instead of developing new platforms.

As an outcome of the analysis, all the identified CCGI platforms that collect data on continuous geographic phenomena do not consider citizen-based QA/QC mechanisms. Geographic phenomena have properties that change much rapidly. Hence, these observations cannot be easily assessed or edited by other users, as long as they cannot be compared to spatial and temporal near observations of known quality. On the contrary, all the VGI platforms and two SGD platforms, which collect data about geographic features have citizen-based QA/QC mechanisms. The existence of QA/QC mechanisms is enabled by the fact that GI about features can easily be reviewed by citizens that either observe them at a later time, or they re-interpreter a representation of them e.g. images of plants.

Citizen-based rating mechanisms have different purposes in VGI and SGD datasets. The rating of VGI observations is mostly referred to the VGI thematic and positional accuracy, while the rating of SGD observation to their attractiveness/ likability. SGD observations are associated with the subjective perception of citizen about features and phenomena. This provides new research opportunities but it also highlights two important issues. First the statistical representatively of the collected data and second the transparency in the SGD production. The opportunity to include perceptions from contributors could also be evidence of a mixing of quantitative and qualitative information that previous research agendas had called for [26].

## 5 Conclusions and future work

With the emergence of new Internet applications and mobile devices with numerous embedded sensors, an increasing number of citizens is enabled to potentially contribute various types of GI. Additionally, Internet-based platforms originally meant for socially-oriented purposes are expected to contain more types of geographical, environmental or geo-referenced information, such as weather-tagged photos and messages.

With the plethora of CCGI sources, the selection of a dataset, that fits the data quality requirements (i.e., fitness for use), is a task not always feasible, due to the absence of information on CCGI dataset's quality. Moreover, an on-demand assessment of the CCGI datasets quality is not always possible when reference data of known quality is not available or accessible. The use of CCGI datasets that have not been evaluated in terms of spatio-temporal accuracy and completeness might result in a partial or erroneous understanding of the environment.

In this paper, we have provided a theoretical framework for the evaluation of GI with special emphases on CCGI. Depending on the requirements of an application or research study, and once the proposed framework is fully developed and validated, one will be able to select the type of GI i.e. VGI, SGD, Private GD or PGI, that match the required dataset quality and cost. Moreover, by reviewing the characteristics of the GI collection activity, of the data contributors, platforms and data collection tools, which are listed in Section 2, one can have an indication of the expected accuracy and the spatial distribution of the collected data.

In future work, we will address the validation of the proposed typology. To this end, we will examine VGI and SGD datasets in order to measure the relation between the purpose of data collection and the quality and the cost of the collected data.

## Acknowledgement

We would like to thank our colleagues and two anonymous reviewers for their valuable comments and suggestions.

## References

[1] V. Antoniou, J. Morley and M. Haklay. Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon. *Geomatica*, 64(1):99–110, 2010.

[2] N. Budhathoki, M. Haklay and Z. Nedovic-budic. Who map in OpenStreetMap and why? *Presentation at the State of the Map conference,* Atlanta, USA, 2010.

[3] M. Craglia, F. Ostermann and L. Spinsanti. Digital Earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, *5*(5):398–416, 2012.

[4] D. G. Delaney, C.D. Sperling, C.S. Adams and B. Leung. Marine invasive species: validation of citizen science and implications for national monitoring networks. *Biological Invasions*, 10(1):117–128, 2007.

[5] S. Elwood, M.F. Goodchild and D.Z. Sui. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers*, 102(3):571–590, 2012.

[6] Foursquare. (2014). Venues Platform. Retrieved March 04, 2014, from https://developer.foursquare.com/overview/venues.htm

[7] S. Fritz, L. See, M. van der Velde, R. A. Nalepa, C. Perger, C. Schill, … M. Obersteiner. Downgrading recent estimates of land available for biofuel production. *Environmental Science & Technology*, 47(3):1688–94, 2013.

[8] Geo-Wiki. (2013). The Geo-Wiki Project. Retrieved September 02, 2013, from http://www.geo-wiki.org/

[9] J.-F. Girres and G. Touya. Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, 14(4):435–459, 2010.

[10] M. F. Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

[11] M. F. Goodchild and L. Li. Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1:110–120, 2012.

[12] Google. (2013). What is Google Map Maker? Retrieved September 05, 2013, from https://support.google.com/mapmaker/answer/157176?hl=en&ref_topic=1093469&rd=1

[13] M. Haklay. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B Planning and Design*, 37(4):682–703, 2010.

[14] M. Haklay, S. Basiouka,V. Antoniou and A. Ather. How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. *Cartographic Journal*, 47(4):315–322, 2010.

[15] J. Howe (2006). Crowdsourcing: A Definition. Crowdsourcing. Retrieved July 26, 2013, from http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html

[16] INaturalist. (2013). iNaturalist homepage. Retrieved October 24, 2013, from http://www.inaturalist.org/

[17] T. Koukoletsos, M. Haklay and C. Ellul. Assessing Data Completeness of VGI through an Automated Matching Procedure for Linear Data. *Transactions in GIS*, 16(4):477–498, 2012.

[18] V. Lee. Volunteer monitoring: a brief history. *The Volunteer Monitor*, 6(1):29–33, 1994.

[19] NBN. (2013). New NBN Record Cleaner Rules now available. Retrieved November 08, 2013, from http://www.nbn.org.uk/News/Latest-news/New-Record-Cleaner-Rules-now-available.aspx

[20] NoiseWatch. (2013). About NoiseWatch platform. Retrieved July 20, 2013, from http://eyeonearth.org/map/NoiseWatch/

[21] P. V. Oort. Spatial data quality: from description to application. *Wageningen Universiteit*. PhD thesis, 2006.

[22] C. J. Parker, A. May and V. Mitchell. Understanding Design with VGI using an Information Relevance Framework. *Transactions in GIS*, 16(4):545–560, 2012.

[23] E. S. Raymond. *The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary* (Revised Ed.). O'Reilly Media, Inc. 2001.

[24] L. See, A. Comber, C. Salk, S. Fritz, M. van der Velde, C. Perger, … M. Obersteiner. Comparing the quality of crowdsourced data contributed by expert and non-experts. *PloS One*, 8(7), 2013.

[25] J. Silvertown. A new dawn for citizen science. *Trends in Ecology & Evolution*, 24(9):467–71, 2009.

[26] R.S. Smith. Participatory Approaches Using Geographic Information ( PAUGI ): Towards a Trans-Atlantic Research Agenda. *In 5th AGILE Conference on Geographic Information Science*, Palma, Spain, 2002.

[27] M. Stephens. Gender and the GeoWeb: divisions in the production of user-generated cartographic information. *GeoJournal*, 78(6):981-996, 2013.

[28] Twitter. (2014). About public and protected Tweets. Retrieved March 04, 2014, from http://support.twitter.com/articles/14016-about-public-and-protected-tweets#

[29] USGS. (2013). Crowd-Sourcing the Nation: Now a National Effort. Retrieved September 09, 2013, from http://www.usgs.gov/newsroom/article.asp?ID=3664#.UihUkD_LI48

[30] WideNoise. (2013). WideNoise Homepage. Retrieved September 03, 2013, from http://cs.everyaware.eu/event/widenoise

[31] P. A. Zandbergen and S.J. Barbeau. Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones. *Journal of Navigation*, 64(03):381–399, 2011.