# VGI data collection for supplementing official land administration systems

Mustafa Hameed[1], David Fairbairn[2], Suzanne Speak[3], Maythm Al-Bakri[4]

[1, 2]School of Civil Engineering & Geoscience, Newcastle University, Newcastle NE1 7RU, UK
[3]School of Architecture, Planning and Landscape, Newcastle University, Newcastle NE1 7RU, UK
[4]School of Surveying, College of Engineering, University of Baghdad, Baghdad, Iraq

18th - 21st April 2017

**Summary**

This article explores the process of VGI collection by assessing the relative usability and accuracy of a range of different methods (Smartphone GPS, Tablet, and analogue maps) for data collection amongst different demographic and educational groups, and in different geographical contexts. Assessments are made of positional accuracy, completeness, and data collectors' experiences with reference to the official cadastral data and the administration system in a case-study region of Iraq. Ownership data was validated by crowd agreement. The result shows that successful VGI projects have access to varying data collection methods.

**KEYWORDS:** Volunteer Geographic Information, Positional accuracy, Land administration systems

## 1. Introduction

The term 'volunteer geographic information' (VGI) was introduced by Goodchild (2007), to describe the widespread participation of the private citizen in creating geographic information, a function that for centuries had been reserved to official agencies (Goodchild 2008). The majority of VGI projects have concentrated on building web applications that allow volunteers, using the internet and contemporary technology, to access and edit or create features with reference to maps, satellite images and ground-based methods. However, such projects may not always be successful in developing countries, such as Iraq, where internet connections may not be good, the majority of citizens have no knowledge of using web map applications, and there are problems in engaging with communities, especially in rural areas. This paper presents a study in VGI data collection, involving a wide range of community members, and several digital and analogue collection methods acceptable to them, to investigate contributions to official land administration systems (LAS). The paper is based on fieldwork in Iraq in 2016.

### 1.1. The potential for VGI in LAS

Only 25% of nations (mostly industrial countries, 35-50 in total) have a complete land registration system: lack of finance, limited institutional capacity and ineffective political will, mean that 75% of the world's land parcels have not yet been registered. The majority of their occupants are the most vulnerable and poorest groups in society and they live in threat of expulsion due to lack of security of tenure (McLaren 2013). It is suggested that cooperation with the local community could accelerate the creation of a 'fit for purpose' system to improve this situation.

The priority of a 'fit-for-purpose' system is not necessarily high spatial accuracy, but rather effective recording of ownership and provision of security of tenure for underprivileged communities (Enemark al. 2014). Building such a system could well incorporate efforts of local volunteers, with different level of education and background, and utilising varying technologies.

[1] m.r.hameed@newcastle.ac.uk
[2] dave.fairbairn@newcastle.ac.uk
[3] s.e.speak@newcastle.ac.uk
[4] m.albakri@coeng.uobaghdad.edu.iq

The majority of previous research in VGI has concentrated on the general drive towards, nature of, and applications for, crowdsourced data. It focussed mainly on the nature of projects such as OpenStreetMap (OSM) and wikimapia, with little research on the use of VGI in official domains, such as land administration. Keenja et al. (2012) reported that "to date, limited empirical work has been undertaken in this domain: there remain many unanswered questions regarding the accuracy, authority, assuredness, availability, and ambiguity of crowdsourced data. Meanwhile, the potential for crowdsourcing to provide a low cost and high-speed solution in areas where cadastral coverage is lacking, is eagerly anticipated" (p2). A more recent project (Basiouka et al., 2015) used volunteers to assess the possibility of using OSM for official, cadastral purposes, but the target group was educated people, perhaps college students, rather than real members of the community being assessed.

In contrast, the research presented here addresses local citizens, variable data collection methods, and also considers whether VGI could be an opportunity or threat to official geospatial systems.

## 2. Methodology of the study

Field work was conducted in the region of Al-Hillah, Iraq. Here, three types of locality - rural, peri-urban and urban - were identified. For each, several specific locations for field work were chosen and the local communities were contacted through their leaders. These different locations are chosen to test which data collection methods work best in varying geographical, topographic and socio-cultural contexts (Figure 1).

Three data collection methods were developed for the local volunteers to use:

1. Smartphone with a GPS app uploaded for locating and attributing land parcel corners;
2. Portable iPad Tablet PC with cadastral map uploaded, and overwriting and annotating capability provided through QGIS;
3. Paper printed aerial image with clipboard and pencil for demarcation.

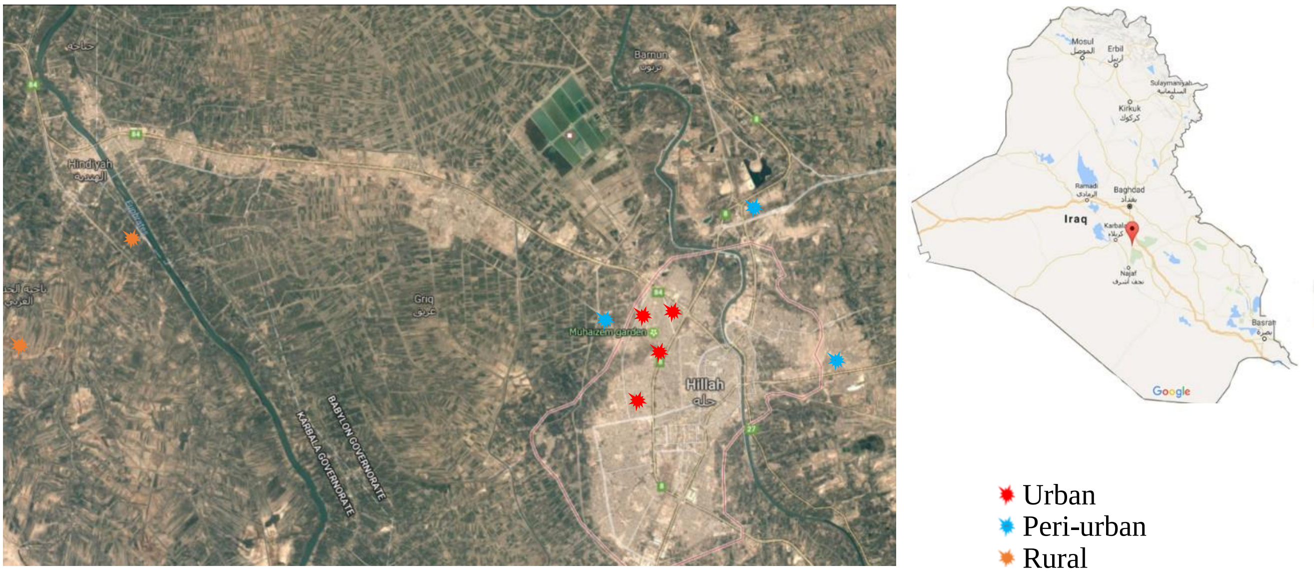


**Figure 1** Study locations, Al-Hillah, Iraq (Google Maps, 2016)

## 3. Activities and outputs

A series of field data collection activities were undertaken in the case study areas. Communities were consulted, volunteers identified, officials interviewed, data collected, and existing official mapping compared. Examples of the main issues in data collection were identified as follows:

### a. Positional accuracy of geospatial data.

Volunteers were given training and written guidelines to enable them to identify the boundary points of their own plot, or others' properties, using the different methods (Figure 2).

**Figure 2** Data collection methods applied in different geographical contexts (Smartphone GPS in urban centre; Tablet computer in urban centre; Paper aerial image in rural area)

This prepared volunteers for the subsequent work of using each method in each of the communities and validated their ability to take accurate measurements. The geospatial data collected by the volunteers was examined and analysed to compare the positional accuracy of the three different methods, in relation to the formal / official data collected by professional surveyors in the Al-Hillah LAS (Table 1).

**Table 1** Root mean square error (RMSE) for parcel corners for compared datasets

| Study area | No. of points tested | RMSE (metres) cf. Official data | | |
|---|---|---|---|---|
| | | Smartphone GPS | iPad Tablet | Analogue paper photo |
| Urban (4 sites) | 778 | 4.364 | 1.357 | 2.615 |
| Peri-urban (3 sites) | 308 | 2.933 | 1.354 | 2.190 |
| Rural (2 sites) | 139 | 3.23 | - | 3.41 |

Main finding – The results were consistent across each site within each category, verifying that using the iPad for spatial data collecting in both urban and peri-urban areas is the best choice in regard to positional accuracy. However, the iPad was not applied in the rural areas, where local people were reluctant to use it. Here, the alternative methods (Smartphone and analogue) were similar to each other.

### b. Completeness of attribute data.

This considers the number and size of plots, completeness, and currency of recorded information. Initially the number of plots has been compared for each case study area, and a significant difference between the formal and volunteer numbers has been found, as follows:

- Urban 1235 plots on the official map = 2133 plots observed by the volunteers
- Peri-Urban 223 plots on the official map = 285 plots observed by the volunteers
- Rural 80 plots on the official map = 728 plots observed by the volunteers

Interpretation – In the urban area there has been significant change due to sub-division of plots; but peri-urban zones exhibit less change because of lower land values (less pressure to sub-divide) and stable occupancy. In rural areas, some parts are very stable, but other sites have changed land use from agriculture to residential, and numerous housing plots have been created from one field.

For example, in the urban case study, one land parcel, believed by the authorities to be a single plot, was shown to have been sub-divided into three separate plots. Two of these were used for housing and the third consisted of several small shops, with a flat above. In another case, one owner of two individual large housing land plots has combined them to produce one plot, now containing five houses. Other locations highlighted a radical change of use, from purely residential to multiple use, including areas where all houses now have shops at the front.

Main finding – based on the attributes of plots – owner, land use, last transaction, and type of tenure - collected from all communities, it is clear that the official register and maps are completely out of date. The Municipality does not have the capacity to update or maintain the system, and earlier attempts to sub-contract improvements to the land administration system to external (Turkish) consultants failed.

### c. Other attribute characteristics

Ownership was validated by crowdsourcing (i.e. multiple data collection) because a) official records of ownership are confidential; b) subdivision has complicated ownership; c) up-to-date information is required.

**Table 2** Verifying ownership data by crowdsource agreement

| Study area | No. of plots tested | No. of plots with inconsistencies in named owner | Percentages of error |
|---|---|---|---|
| Urban (4 sites) | 200 | 9 | 5% |
| Peri-urban (3 sites) | 150 | 5 | 3% |
| Rural (2 sites) | 80 | 2 | 2% |

Main finding – The percentage of disagreement in validating ownership data, by consensus, was low across the three different zones. It is concluded that ownership data obtained from groups of individuals were correct.

### d. Experiences in engaging with volunteers

Other factors considered during the fieldwork included the preferences and work done by the volunteers, analysed from interviews, observation and analysis of the data captured. Feedback sessions have been held with volunteers in the peri-urban community after their participation, showing the resulting maps of their work and recording their opinions.

Main finding – Citizen volunteers have valuable information on their communities that, if added to the formal land administration system, can help to update it. For example, in one of the case study areas, volunteers have reported that the official map is old because a river that is shown on the map no longer flows and the area of land is occupied by some residents. Another example relates to permissions: the out-of-date official map shows a block of land parcels which appear to be available for development, but in reality the area of that block is now considered a heritage area.

## 4. Discussion

Analysis has shown the relative accuracy of different data collection methods in different contexts, and future work will explore why some methods were more successful or accurate. We also argue that, in some cases, it may be more important to collect some interim data, which the community can agree on and take ownership of, even if that means using a slightly less accurate method, than to focus simply on spatial accuracy. It can be concluded that in areas of conflict, or when official systems are under extreme stress, VGI may be the only realistic method of collecting data. In these cases, it may be more important to encourage people to use a method of their preference and ability, and to sacrifice some degree of spatial accuracy.

## 5. Conclusion

This paper has concentrated on considering different methods of data collection that can suit different types of people, in varying geographical contexts. The research has given opportunity for the majority of volunteers to participate, no matter what their level of education or experience is. The encouraging levels of accuracy and completeness of the VGI data, have interested the land administration system authorities of Al-Hillah.

**References**


Basiouka, S. and Potsiou, C. (2015). "OpenStreetMap for cadastral purposes: an application using VGI for official processes in urban areas." Survey Review 47(344): 333-341.

Enemark, S., Clifford Bell, K., Lemmen, C. and McLaren, R. (2014). Fit-for-purpose land administration. Copenhagen, Denmark, International Federation of Surveyors (FIG).

Goodchild, M. F. (2007). "Citizens as sensors: the world of volunteered geography." GeoJournal 69(4): 211-221.

Goodchild, M. F. (2008). "Commentary: whither VGI?" GeoJournal 72(3-4): 239-244.

Keenja, E., De Vries, W., Bennett, R., & Laarakker, P. (2012). Crowd Sourcing for Land Administration: Perceptions within Netherlands Kadaster. In Proceedings of FIG Working Week, Rome (paper 5611).

McLaren, R. (2013). Engaging the Land Sector Gatekeepers in Crowdsourced Land Administration. Proceedings of World Bank Land and Poverty Conference, Washington.