

ROADPLEX: A MOBILE VGI GAME TO COLLECT AND VALIDATE DATA FOR POIS

Alireza Kashian^{a*}, Abbas Rajabifard^a, Kai-Florian Richter^b

^aDepartment of Infrastructure Engineering, the University of Melbourne, Australia –
a.kashian@student.unimelb.edu.au, abbas.r@unimelb.edu.au

^bDepartment of Geography, University of Zurich, Switzerland – kai-florian.richter@geo.uzh.ch

KEY WORDS: VGI, Points of Interest, Gamification, Games with purpose, RoadPlex

ABSTRACT:

Because of increasing popularity of smart phones equipped with GPS sensors, more volunteers are expected to join VGI (Volunteered Geographic Information) activities and therefore more positional data will be collected in shorter time. Current statistics from open databases such as OpenStreetMap reveal that although there has been an exponential growth in the number of contributed POIs (Points of Interest), the lack of detailed attribute information is immediately visible. The process of adding attribute information to VGI databases is usually considered as a boring task and it is believed that contributors do not experience a similar level of satisfaction when they add such detailed information compared to tasks like adding new roads or copying building boundaries from satellite imageries.

In other crowdsourcing projects, different approaches are taken for engaging contributors in problem solving by ‘hiding’ the tasks behind a game. In the literature, this concept is known as “gamification” or “games with purpose” which encapsulate the idea of entertaining contributors while they are completing a particular defined task. The same concept is used to design a mobile application called “RoadPlex” which aims to collect attribute information for POIs.

The increased number of contributions in the past few months confirms that the design characteristics and the methodology of the game are appealing to players. Such growth enables us to evaluate the quality of the generated data through mining the database of answered questions. This paper discusses some contribution results and emphasises the importance of using gamification concepts in the domain of VGI.

1. INTRODUCTION

In recent years, advances in positioning, web mapping and communication technologies have changed the traditional methods of collecting, updating and maintaining geographic information. People distributed within an environment can voluntarily participate in collaborative online activities and help producing knowledge about any object that has geographic attributes. Accordingly, various desktop and mobile applications have been designed that are dedicated to empower the crowd for missions of data collection. There is no doubt that the results of such collective work are valuable and can be beneficial for both citizens and governments. While each person works on a small task, the final collective result often turns out greater than its parts. The term of VGI (Volunteered Geographic Information) was coined by Michael Goodchild and highlights the potential use of "Citizens as voluntary sensors" to create and enhance geographic data (Goodchild, 2007). VGI encapsulates the process of collecting, contributing, maintaining, and processing of geographic data (short: geodata) by volunteers. In the past decade, OpenStreetMap, Wikimapia and Google Map Maker have proven this process as a successful phenomenon that is unlikely to go away again. However, while having the opportunity of getting access to free geodata sounds appealing, the quality of the generated data is debatable and needs further investigation. This could be one reason why most governmental and professional agencies who are involved in processing and maintaining geographic information are encouraged to only deal with authoritative data which is collected and processed by

professional and trusted sources, under their own supervision. Often authoritative data is the only available resource for legal dispute, for example, for claims such as land ownership and political arguments about administrative boundaries.

We are interested in studying fundamental behaviours and properties of VGI platforms. Thus, we designed a new mobile VGI game called RoadPlex in order to investigate new methods of collecting and examining the validity of geodata collected by ‘the crowd’. RoadPlex can be considered a platform since it provides a systematic approach towards collecting and managing geodata. This new VGI platform opens up opportunities to investigate some interesting topics, such as understanding motivations of contributors (Budhathoki 2010) and improving techniques to increase the speed of data collection and expression of how good VGI data is (Haklay 2010). Throughout this paper, we discuss the importance of points of interest and highlight how RoadPlex can help to collect and validate description data about point features by employing the gamification concept.

2. CROWDSOURCING POINTS OF INTEREST

VGI as a source of geodata has been attractive to academia as well as businesses and governmental agencies. It seems like a myth that the world is well mapped, and maps are getting updated regularly in order to provide better accuracy and completeness (Mooneyhan, & Estes, 1994). In many parts of the world even critical information about location of emergency services, such as hospitals, or basic road maps are not available, which is vital, for example, in the case of natural

* Corresponding author

disasters. We might think that no extra tasks remain for volunteers to accomplish as the current state of spatial data is satisfactory especially in developed regions of the world, but there are on-going challenges for researchers and businesses alike to verify the accuracy of geodata and to complete attributes and update outdated data. Most VGI initiatives have embedded solutions for collecting and managing point features that refer to a specific spot on the map (or on earth). These locations are called points of interest (POI). There is a wide range of geographic objects that may be considered POIs. For example, post-boxes may not be what comes immediately to mind, but for specific users and tasks they may well be important and, thus, POIs. The concept POI is used differently in different GIS contexts. A proper definition is not given here as the exact nature of POIs is not important for the introduced application. The following list provides some examples of what may be considered POIs in general:

- Churches, hospitals, schools, town halls, distinctive buildings
- Post offices, shops, post-boxes, telephone boxes
- Pubs (pub names are useful when navigating by map)
- Car parks and lay-bys (and whether free or not)
- Bus stations, metro stations, ferry terminals, airport terminals
- Speed cameras, police stations
- Tourist attractions

The rise of cheap mobile devices equipped with GPS sensors and the emergence of Web 2.0 have made it easy for ordinary contributors to view maps, determine the validity of data and share their own spatial experience. Potentially, there are seven billion humans who all possess some knowledge about certain properties of the earth's surface (Goodchild, 2007). If we design proper tools to help everyone at any level of knowledge to participate in the process of creating geographic information, we would be able to get access to this incredible wealth of local information. While getting every human being on the world to participate is obviously an unrealistic goal, this great potential has still inspired us to implement RoadPlex to at least start tapping into this wealth of knowledge. In designing RoadPlex we focus mainly on positional accuracy and attribute completeness for POIs. We decided to select one of the publicly accessible POI data sources to start our work. Among the available popular sources, such as OpenStreetMap, Yelp, Google and FourSquare, we chose FourSquare. There were three basic reasons for this selection:

1. FourSquare itself is considered a crowdsourcing platform where people can insert new POIs and share their own experience.
2. Due to the inherent lack of accuracy and the incompleteness of attributes in FourSquare, there is a real need for our development.
3. FourSquare provides an easy API interface for making robust queries and leading to fast implementation.

FourSquare is a social networking platform mainly developed for smart phones. Users of FourSquare can 'check in' at venues using a device-specific application for GPS-enabled mobile devices. FourSquare provides a list of nearby venues to ease choosing the right venue. Users can use the website or a simple text message to check in at venues. For each venue (which may

be called POI), FourSquare keeps a record of a limited number of attributes, such as name and address, to make the job as simple as possible for users. Selection of these attributes has been the sole decision of FourSquare management, and there is no published statement from FourSquare explaining their methodology for selecting such attributes. The complete list of used attributes (FourSquare, 2013) is listed in Table 1. In the remainder of this paper, we will explain how FourSquare venues are fed into RoadPlex in order to check their positional accuracy and to add more attribute data for each venue.

Key	Description
street address	street name and house number
Locality	the city the venue is in
Region	the administrative region the venue is in
ISO3	the country code
postal code	in the preferred local format for the venue
Phone	in the preferred local format for the venue, including long distance code
Twitter	twitter screen-name, without @
url	official/canonical URL for venue
app:id	An ID or comma separated list of IDs representing the venue in the database

Table 1. Attributes used by FourSquare to describe a venue

3. DESIGN OF ROADPLEX

The collection of VGI data can be categorized into three major methods: geo-tagging available information; collecting GPS-trajectories; annotating satellite imagery. The success of VGI depends on the number of active users ready to contribute. When investigating in methods of encouraging stronger participation, it is important to have users equipped with proper tools matching their expertise and technical skills. Otherwise if the tools are too complex, potential contributors might refrain from continuing to use the tools. And if users do not see the value of their contribution, their incentive to put more effort in contributing will diminish. Taken these factors into consideration, RoadPlex is designed as a mobile VGI game with the hope of lowering the barrier for data contribution (cf. Schmid et al. 2013). Other similar approaches are the Kort application (Hunziker et al, 2012) and the ESP game (von Ahn 2006).

In the field of gamification of crowdsourcing tasks, RoadPlex concentrates on an easy user interface and a straightforward rewarding scheme, so all contributors can immediately see and understand the outcome of their efforts. At the time of writing, RoadPlex supports only Android devices. The latest version of RoadPlex (2.5.3) is available through the Google Play Store under the category of "Travel and Local"².

3.1 Scenario

Once the application is installed, the user needs to register and log into the application. The registration is required as we need to keep track of activities of each player. The application needs to have Internet access via mobile network or WiFi. It also asks the user to enable the GPS sensor. In the beginning, RoadPlex collects the user's positional data from the GPS sensor and sends this data back to the server. The server queries FourSquare through API calls, requesting data for all nearby

² <https://play.google.com/store/apps/details?id=com.rojina.roadplex>

venues in 1 km radius (the default). FourSquare only provides a maximum of 50 venues, sorted by its own ordering algorithm. The RoadPlex server dispatches the list of venues plus an attached group of questions for each venue. These questions are specifically generated for each type of venue.

At this stage, the user can navigate through a list of venues and see the position of each venue on the map. The task of the players is to select one of the venues they know (reasonably) well and then answer some simple questions about this venue. To accomplish this task, players are expected to have enough local knowledge to be able to respond with correct answers. Depending on the type of venue, the type and number of questions will be different. There are three answer styles: radio items, check box list and text box entry. For example, if the venue is categorized by FourSquare as a restaurant, the application might ask for food delivery options as a radio list, but if the venue were categorized as a hotel, the application would ask about the name of the owner through a textbox entry. The list of possible questions and the method for generating questions is discussed in Section 3.4.

Once a player submits an answer for each question, bonus points will be rewarded, which are called "Activity Points" in the game. The application sends all answers back to the server for storing purposes. RoadPlex tries to conflate answers from two different players, and if it finds a high similarity between both submitted answers, it rewards both players with virtual gold coins. Players who collect enough coins will receive actual prizes in form of Google Play credits. This credit approach will encourage players to submit accurate information in order to have a higher chance of getting matched with other players. The list of high score players is published in the 'Top Score' section of the application, to encourage players to compete with each other to stay at the top of the list.

3.2 System Design

RoadPlex follows a simple client-server logic. The client runs on users' devices as an Android application, while the server is responsible for managing logins, controlling users, sending a list of questions to the client and preparing queries to FourSquare. Clients work independently, and users have no means of communication with each other. The diagram in Figure 1 depicts how clients, server and FourSquare are interconnected.

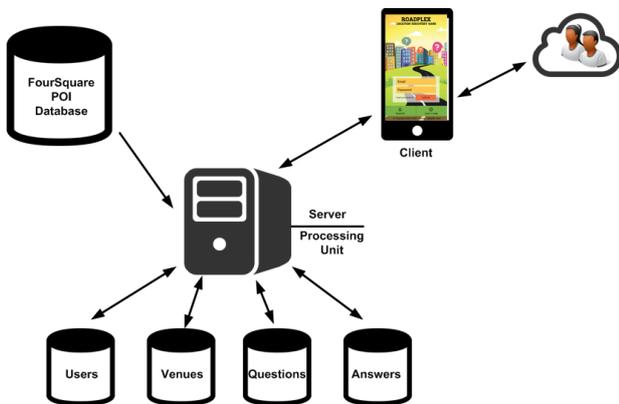


Figure 1. RoadPlex data interactivity diagram

3.3 User Interface

Figure 2 illustrates the user interface of RoadPlex in different stages:

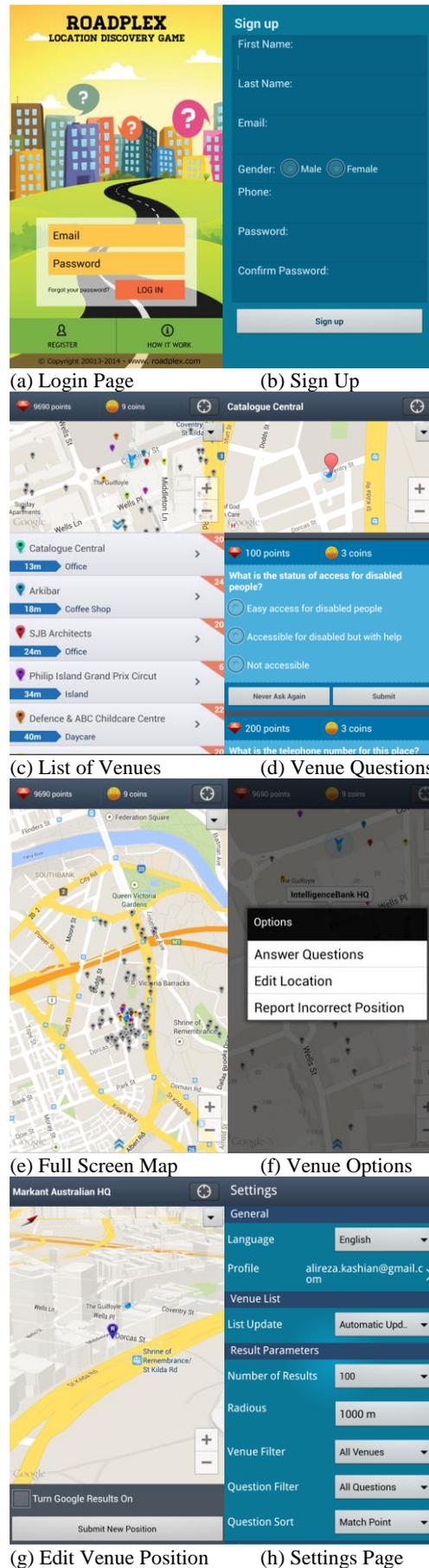


Figure 2. RoadPlex User Interface

(a) Login Page: Through this page, the user can register or log in to the application. Users also have access to a short introduction on how RoadPlex works.

(b) Sign up page: The user needs to enter personal details such as name and email address to sign up for RoadPlex.

(c) List of Venues: The application provides a list of venues. Each venue is represented by a marker on the map, at the top of the screen. There is a small orange triangle with a number at the top right corner of each venue item. This number shows the total number of questions available for this venue. There is also a blue arrow, which indicates the distance of the venue from the current position of the user. Users can scroll through this list. Clicking (a single touch) on a venue will pop up a new page with the list of questions for that venue.

(d) Venue Questions: On this page, users can navigate through a list of questions. Users can submit the answer by clicking on one of the options or by entering data in a text box. Users can also select a "Never Ask Again" button to indicate that they will be no longer interested in answering this particular question.

(e) By clicking on the small blue arrow at the bottom of the map, the map gets expanded to full screen, making it easier to pan and zoom.

(f) Touching any of the markers (venues) on the map provides a list of options such as answering questions for that venue, or editing the location of the venue. Users may also simply report the incorrect position without proposing a new position.

(g) If the user selects the "Edit Location" option, a new full screen map is shown so the user can touch and keep the finger on the marker and then drag the marker to the correct position. The user needs to zoom in adequately to get a detailed view of the area. Users can also select the check box for Google results, to see what Google proposes as the correct location of the venue, and if they agree with Google, they can relocate the marker to the proposed position.

(h) The setting page of RoadPlex allows users to manually adjust a few options. Users can: 1) change the language of the interface; 2) edit profile information such as email address, password and name 3) define the frequency of updating the venue list; 4) set the maximum number of venues in the list; 5) set the default scanning radius; 6) filter the venues for all or only the venues which other users already answered questions for; 7) filter the questions for all questions or those questions which are already answered and waiting to be matched; 8) sort the questions based on match points or latest answered questions.

3.4 Question Structure

As we discussed earlier, players need to answer simple questions for each venue, but how are these questions generated? And what kind of information is useful to be collected?

RoadPlex' methodology of data collection is based on asking easy and quick questions intelligently. By saying "intelligently", we mean that questions should be directly related to the type of venue. So if the venue is a hospital, the

questions might be related to the number of beds available in this hospital, different available treatments or whether the hospital has emergency facilities. To achieve this goal we designed an online administration panel to assign each question to a number of categories, which are extracted from FourSquare's categorization schema. We decided to collect a broad range of attributes. Some generic attributes are: availability of a car park, ATM access, WiFi status or product delivery options. We designed simple questions to ask for the right value of these attributes. The answers are in the form of check box, radio button or, for some questions, in the form of a text box. In the case of checkbox and radio button, we did our best to cover all possible answers and to avoid overlapping options. Two sample questions and answers are:

What type of credit cards does this place accept?

- (a) All three majors Visa/Master/American Express
- (b) Only Visa/Master
- (c) No credit card is accepted

What is the smoking status in this place?

- (a) Smoking is not permitted
- (b) Smoking outside only
- (c) Only in smoking zones

The question list is dynamically evolving and valuable information is captured through the interviews with players who provide feedback, such as whether a particular question is very easy or not easy at all to answer, or whether answers do not have the options they expected.

Current questions include the following:

- What is the status of nearby cashiers / bank?
- What are the possibilities to bring pets into this place?
- What delivery options are available in this place?
- What is the structural status of this place?
- What is the status of access for disabled people?
- What type of credit cards does this place accept?
- What is the status of WiFi Internet here?
- Who is the owner of this place?
- What is the smoking status in this place?
- What is the importance of this place for public visitors?
- What year was this place established?
- What is the status of outdoor/indoor access?
- What is the reservation status for this place?
- What is the website address for this place?
- Which days of the week is this place closed?
- What time does this place close most of the days during the week?
- What time does this place open most of the days during the week?
- What is working hour shifts for this place?
- How big is this place to accommodate people inside?
- What is the parking status for this place?
- What is the email address for this place?
- What is the telephone number for this place?
- What is the address for this place?
- What is the exact name for this place?

3.5 Rewarding Points

Each player receives rewarding points based on each answer they submit for a question. As players need to get actively involved, these points are referred to as "Activity Points". Depending on the level of the effort needed to answer a question, more points are rewarded for some questions. There are four different predefined levels of points: 50, 100, 150, and 200. For example, if a user replies to a question like: "Does this

place have indoor or outdoor space?” players would receive 100 points while a question like: “What is the email address for this place?” gets 200 points. This means users would probably need to investigate more to find the answer for the email address and dedicate more time to enter it.

There is a simple definition for rewarding points. If players need to walk around, seek for help, read boards or talk to people to find an answer, they deserve more points compared to the situation where players can easily find the answer with just a quick look around. Currently the number of questions is limited, so it was easy to assign points to each question manually, but for a more comprehensive list of questions in the future, a more sophisticated pattern to assign points to each question is needed.

3.6 Matching and Validation

If two players provide the same answer to a question for a particular venue, a “double match” event will happen and both players will receive gold coins. Such an event means both players agree on the same answer for the same question about the same venue, which fundamentally may be considered a verified result. These matches will help to tag answers as verified in the venue database. In this case, other players are no longer asked to answer this question, assuming it is already solved. Designing proper questions with a useful list of options will make the probability of getting a match randomly very low, but this probability still needs to be minimized through different methods and considerations. There are three main scenarios when two players match with each other:

- Both players know each other and agreed to cheat. They agreed on providing the same (not necessarily correct) answers to the questions of a venue.
- Both players coincidentally provide the wrong answer to a question. Depending on the type of question, chances for this are very low (text box) or reasonably probable (radio buttons).
- Players faithfully reply with the correct answer. Both answers are true, and both players deserve to get the coins.

While RoadPlex passes through early stages of live tests, we currently collect answers and try to understand patterns of replies. To overcome some of the especially fast random replies and collusions between players, a simple algorithm is developed to identify those patterns. If a reply from a player arrives very soon after seeing the question, it could be assumed that this player did not spend enough time to provide correct answers, especially for those questions that require some local investigation. If two players are detected to constantly getting matched, it will be suspicious as there is a reasonable assumption that players do not know each other, and the chance of two random players getting matched frequently in a short period of time is very low. In this case, we tag both players as suspicious players and deal with their answers carefully.

3.7 Contribution Result

By February 2014, RoadPlex attracted around 732 registered players with approximately 5000 questions answered with a gradually growing number of participants. There are around 37 distinct questions in the system and 287 players have answered at least one question.

Through our analysis, we found that some players had responded to some questions randomly in order to receive more points in shorter time. While this is not a prevailing behaviour among the players, the consequences of such activities decrease

the reliability of responses. To control such behaviour, a double blind matching algorithm is needed to attest the validity of each answer by asking two anonymous players the same question. It is also planned to run a competition inside the campus of the University of Melbourne, which would help to test our blind matching algorithm with a smaller, but more controlled community.

Some of the latest results on top ten players based on received points, and top questions receiving most answers are shown in Tables 2 and 3. It is also interesting to understand which locations the top 5 players had the most concentration on. Figure 3 shows that except the player ranked first the other players have focused on smaller regions where they probably live or work, which indicates they have more attention on their local neighbourhood.

Player	Number of Answered Questions	Received Points (Descending)
Player 1	405	54100
Player 2	352	49100
Player 3	278	45130
Player 4	258	42340
Player 5	177	28590
Player 6	159	22000
Player 7	157	21900
Player 8	163	21810
Player 9	129	20270
Player 10	124	19350

Table 2. Top ten players based on received points.

Question	Number of Answers
Is the position of the venue accurate on the map?	730
How many people would be able to get into this venue at the same time?	687
Does this venue provide free WiFi for the public?	682
Is this venue located indoor or outdoor?	680
What are the possibilities to bring pets into this place?	181
Which facilities does this place have?	159
What is the location address?	150
What is the importance of this place for public visitors?	149
What is the structural status for of this place?	144

Table 3. Top questions receiving most answers.

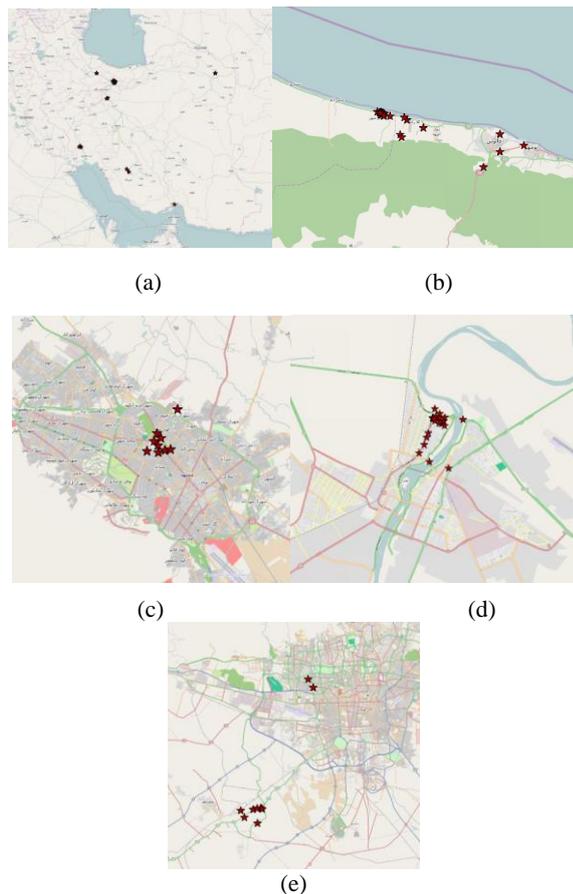


Figure 3. Propagation of POIs for the top ranked player (a), rank 2 (b), rank 3 (c), rank 4 (d), and rank 5 (e).

4. CONCLUSION

In this paper we introduced RoadPlex as a new VGI platform to collect and verify data related to attributes of points of interest (POIs). We explained the mechanisms and the user interface of RoadPlex to highlight its important features. We discussed the importance of POIs in today's online mapping engines and motivated our intention to fill the gaps of description information with a gamification approach. RoadPlex prepares the grounds for collecting valuable pieces of information through asking simple questions about different POIs. These POIs are collected from FourSquare's database which itself is considered a VGI application.

The RoadPlex game provides an opportunity to identify patterns of malicious entries and to verify inputs through matching answers between two players. While RoadPlex is still at early stages, there is good opportunity to gain more useful results through inviting more players to join. The recent logs show that some players already have more than 200 answered questions in their profile which demonstrates that the idea of this platform is interesting enough for some users to get actively involved in the process of generating data for longer periods of time.

Future work is geared toward upgrading the user interface, changing the scenarios for collecting data, establishing quality metrics of the submitted data, and enhancing the verification and matching algorithms. Through these enhancements, we expect to generate higher quality data that may be useful for a wide range of applications especially for authoritative and official purposes.

REFERENCES

Budhathoki, NAMA R A J. 2010. Participants 'motivations to contribute geographic information in an online community, Urban & Regional Planning, University College London.

FourSquare, 2014. About Geo Place Attributes. <https://dev.twitter.com/docs/> (20 Feb 2014)

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

Haklay, Mordechai. 2010. "How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets." *Environment and Planning B: Planning and Design* no. 37:682-703. doi: 10.1068/b35097.

Hunziker, Jürg, and Stefan Oderbolz. Gamified Mobile App für die Verbesserung von OpenStreetMap. Diss. HSR Hochschule für Technik Rapperswil, 2012.

Mooneyhan, D.W., Estes, J.E., 1994. Of maps and myths. *Photogrammetric Engineering and Remote Sensing*, 60(5), pp.517-524.

Schmid, Falko, Lutz Frommberger, Chunyuan Cai, and Frank Dylla. 2013. Lowering the Barrier How the What-You-See-Is-What-You-Map Paradigm Enables People to Contribute Volunteered Geographic Information. Paper read at Proceedings of the 4th Annual Symposium on Computing for Development - ACM DEV-4 '13, at New York, New York, USA.

Von Ahn, Luis. "Games with a purpose." *Computer* 39.6 (2006): 92-94.