

# Evaluating the VGI Users' Level of Expertise: An Application of Statistical and Artificial Neural Network Approaches

Elaheh Azariasgari, Islamic Azad University, South Tehran, Iran

Farhad Hosseinali, Shahid Rajaee Teacher Training University, Iran\*

## ABSTRACT

Currently, online maps are of the most innovative and significant sources of information in people's daily life. However, quality assessment of volunteered geographic information (VGI) data raises some challenges. This research aims to analyze the VGI participants' level of expertise through evaluation of their background information. Towards this goal, an android application was developed to test users' knowledge and cognition about some selected regions of city as well as their background information. In order to evaluate the quality of information expressed by participants, some changes were made in Tehran's online map, and users were asked to identify the changes and to guess the vanished attributes. Statistical and ANN approaches were applied for analysis. The results demonstrated that the ANN was able to predict the percentage of correct answers of a new volunteer with mean squared error of 0.2. This research suggests that users' age and familiarity with the specific region in the city play more significant roles in their expertise in using online maps and in probable participating in VGI.

## KEYWORDS

Artificial Neural Network, Data Quality, GIS, Statistical Analysis, Tehran, VGI

## INTRODUCTION

In recent years, there has been a tendency to use the Web to create, assemble, and disseminate geographic information provided voluntarily by individuals. The users are also interested in contributing to the field of geographic information called volunteered geographic information (VGI) (Goodchild, 2007). VGI has the potential to be a significant source of geographic information at the Earth's surface. However, there are many challenges with VGI (Kessler, 2011). VGI is based on geographic maps and also is a progressive field of study in the former science and technology of geographic information systems (GIS) (Gómez-Barrón et al., 2019). Urban rapid expansion and changes also make it difficult to provide formal up to date maps for everyday works (Hosseinali et al., 2014; Lin, 2018). Websites such as Wikimapia and OpenStreetMap are samples of VGI-based systems that often provide the cheapest source and sometimes the only source of geographic information (Ali

DOI: 10.4018/IJAGR.316770

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

& Schmid, 2014; Da Costa, 2016). However, VGI data can be even obtained from the websites whose activities are in other areas and allow users to add geographic information as well (Astaburuaga et al., 2022; Feng & Sester, 2018).

Nevertheless, integration of VGI with valid data, Spatial Data Infrastructure (SDI), and other data has emerged as an issue, which is due to the fact that it is difficult to accurately determine the quality of VGI (Fogliaroni et al., 2018; Severinsen et al., 2019). There are two main methods to assess VGI data quality. In the first approach, it is assumed that formal authoritative data is quality data and can be used for VGI data validation. The most important disadvantage of this method is that it requires access to formal authoritative data, which might not be always possible. Also, correctness and consistency of authoritative data cannot be guaranteed, which is due to the slow update rate of authoritative data-sets, resulting in discrepancies between data and reality that arise over time. In the second approach, the evolution of data, namely its history or progress is analysed. The advantage of this approach compared with the first one is that it does not require external data sources and considers the timeliness dimension of data. However, this method does not measure data quality accurately, but instead provides an approximation (Antoniou & Skopeliti, 2015; Dai et al., 2008; Severinsen et al., 2019).

Research has shown that three indicators of practice, skill and motivation can affect the quality of the data produced (Yang et al., 2016). Also, the concept of informational trust has been introduced to the established notion of interpersonal trust, and it has been suggested to use informational trust and reputation as proxy measures for information quality. The question that is raised here is that whether users' background information including occupation, age, etc. can be an indicator of people's level of expertise in producing volunteered geographic information? And if this information plays a role in validity of produced data? Therefore, in this paper, we aim to evaluate the quality of information expressed by VGI participants (in an urban environment) based on their background information, by using Artificial Neural Network (ANN) as well as statistical approaches. The next section discusses related work and illustrates the necessity of our work. Then, the methodology is explained, and the results are expressed and briefly discussed. Finally, the last section summarizes the article and provides recommendations for future works.

## BACKGROUND

This part of the study reviews different methods and approaches that have been used by previous researchers to assess the VGI data quality.

Goodchild and Li (2012) described three approaches to quality assurance, which were termed the crowd-sourcing, social, and geographic approaches. Crowd-sourcing approaches "refers to the ability of a group to validate and correct the errors that an individual might make" while social approaches "relies on a hierarchy of trusted individuals who act as moderators or gate-keepers". The geographic approach is powered by certain rules of syntax that reveals what can and cannot occur at a given location (Goodchild & Li, 2012). In another work, Antoniou and Skopeliti (2015) addressed the various aspects of measuring VGI's quality (Antoniou & Skopeliti, 2015). Senaratne et al. (2017) reviewed the literature of VGI data quality assessment and classified them into three categories of: map-based, image-based and text-based (Senaratne et al., 2017). Bordogna et al. (2014) addressed the ways of increasing VGI data quality (Bordogna et al., 2014). Moreri et al. (2018) proposed a Trust and Reputation Modelling (TRM) methodology to measure the quality of VGI data in Mochudi, Botswana without the typical reference to ground truth. TRM measurements involved: Thematic accuracy measure, Semantic accuracy measure, Volunteer credibility determination and Positional accuracy determination. They underlined that "metadata about contributed data sets can be created by the development of rating applications for the public to police themselves in assessing and subjectively rating the accuracy of other volunteer contributions" (Moreri et al., 2018).

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/evaluating-the-vgi-users-level-of-expertise/316770](http://www.igi-global.com/article/evaluating-the-vgi-users-level-of-expertise/316770)

## Related Content

---

### Simulation-Based Total Energy Demand Estimation of Buildings using Semantic 3D City Models

Robert Kadenand Thomas H. Kolbe (2014). *International Journal of 3-D Information Modeling* (pp. 35-53).

[www.irma-international.org/article/simulation-based-total-energy-demand-estimation-of-buildings-using-semantic-3d-city-models/120064](http://www.irma-international.org/article/simulation-based-total-energy-demand-estimation-of-buildings-using-semantic-3d-city-models/120064)

### Between Cuvier and Darwin: Issues in a Changing Heritage

Stefano Brusaporci (2019). *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 1730-1744).

[www.irma-international.org/chapter/between-cuvier-and-darwin/222970](http://www.irma-international.org/chapter/between-cuvier-and-darwin/222970)

### An Integrated Application Ground Penetrating Radar and Seismic Refraction for Non-Intrusive Investigation of Geophysical and Geotechnical Targets

Kebabonye Laletsangand Lucky Moffat (2018). *Handbook of Research on Geospatial Science and Technologies* (pp. 138-149).

[www.irma-international.org/chapter/an-integrated-application-ground-penetrating-radar-and-seismic-refraction-for-non-intrusive-investigation-of-geophysical-and-geotechnical-targets/187723](http://www.irma-international.org/chapter/an-integrated-application-ground-penetrating-radar-and-seismic-refraction-for-non-intrusive-investigation-of-geophysical-and-geotechnical-targets/187723)

### Toward an Architecture for Enhancing Semantic Interoperability Based on Enrichment of Geospatial Data Semantics

Mohamed Bakillahand Mir Abolfazl Mostafavi (2012). *Universal Ontology of Geographic Space: Semantic Enrichment for Spatial Data* (pp. 53-72).

[www.irma-international.org/chapter/toward-architecture-enhancing-semantic-interoperability/63995](http://www.irma-international.org/chapter/toward-architecture-enhancing-semantic-interoperability/63995)

### Control of Full-Size Frame Structures via Optimum Tuned Mass Dampers

Apaer Mubuli, Sinan Melih Nigdeliand Gebrail Bekda (2021). *International Journal of Digital Innovation in the Built Environment* (pp. 47-61).

[www.irma-international.org/article/control-of-full-size-frame-structures-via-optimum-tuned-mass-dampers/277121](http://www.irma-international.org/article/control-of-full-size-frame-structures-via-optimum-tuned-mass-dampers/277121)