Emerging Technology for Data Collection and Citizen Science

[1]Vinay Gangare and [2]Mridul Das*
[1]Researcher, Department of Earth Science, Barkatullah University, Bhopal, Madhya Pradesh - 462026
[2]Assistant Professor (Contractual), Department of Geography, Serampore College, Serampore, Hooghly, West Bengal – 712201

Corresponding Author:
Dr. Mridul Das

ABSTRACT: The daily data collection from a human being becomes the new trends in the day to day life. This activity engages local citizen to generate empirical information. Crowdsourcing or data gathering is a technique utilized for acquiring qualitative information through web technology. Through the use of Geospatial web-based technique, this qualitative information plays a crucial role to generate a decision-based platform. We performed a citizen-based data acquisition experiment to produce a web-based platform for an information system, all the information collected from the citizen identified and verified with available online and offline resources which gathered into a single web-based platform. The open source technology ODK collect and QGIS data collection, and spatial visualization tool applied for data collection experiment. Finally, a website was built to provide integrated information about the place for tourists as well as commoners.

KEYWORDS: Geospatial, ODK collect, Citizen science, QGIS

INTRODUCTION

Geospatial technology is the specific term used to describe the variety of modern tools (thematic mapping, telemetry, Remote sensing, GIS, and satellite positioning systems such as the Global Positioning System) contributing to the geographic mapping and analysis of the earth and human societies. Geospatial technology integrated with information system helps users, researchers and, organizations to generate, collect, store, modify and, analyze the data. Geo-Enable Information System powered by citizen science is a system to integrate community-driven research or citizen science with the Geo-Enable Information System. Whereas, citizen science is the involvement of the public in scientific research, whether in community-driven research or global investigations. In scientific research, citizen science discussed as public science or public participation, and it is a rising movement that accounts the public in scientific discovery, monitoring, and experimentation over an extensive range of disciplines. Different scientific branches in the world, i.e., microbiology (Cooper et al., 2010) to astronomy (Lintott et al., 2008), are using this volunteer participation. The success of these citizen science projects headed to better collaboration between the participants and scientists, and citizen scientists. Now a day this science discovered the mysterious characteristics of hundreds of thousands of galaxies from the Sloan Digital Sky Survey (Lintott et al., 2008), estimated protein structures using the fold it game (Seth, C. et al., 2010), and provided improved solutions to the Multiple Sequence Alignment problem of comparative genomics (Kawrykow, A. et al., 2012).

Tourism is one of the prominent sources for the socio-economic development of any region. Moreover, this factor is considered as the prime aspect in the case of hill stations. Therefore, it is crucial to monitor the overall condition of a hill station from time to time, for not only the betterment of the tourist destination but also to attract people by its services. A survey by using citizen science is the best possible method for active and rigorous monitoring of such a dynamic place. The present report takes Point of Interest (POIs) for tourists' amenities and services as a case to demonstrate the application of Geo-Enable information system using citizen science. Mussoorie (The Queen of Hills), was taken as a case study for the research. It is a wonderful hill station in the Uttarakhand state of India, situated at the height of 2005 m above the sea level. In this research, we analyzed both resident and tourist participants of the citizen science program, and incorporate the survey outcomes to discourse our research questions and achieve a better understanding of variations among inhabitants, and tourist volunteers. It is hypothesized that a volunteer's place affection, the emotional association between a person and place, can affect their stimuli to participate in the project (Ellis and Waterton 2004; Budruk and White 2008; Rotman et al. 2012; Nov, Arazy, and Anderson 2014). However, volunteer's residency status is anonymous, either resident or tourist's place attachment impacts the data quality and

ultimately effects volunteer and scientific outcomes of a project. This project hypothesizes that after participating in a citizen science project, tourists can evoke place attachment at an equal level, or better than residents. The incorporating technology of mobile phone-based applications, a tourist-centric citizen science program will produce high quality educational and scientific outcomes, which are similar or better than the resident-participants in the same program.

METHODOLOGY

The methodology of the research was divided into four parts (fig - 1). The first part of the methodology includes the assignment of the study area, the process of problem identification, followed by data collection, filtration of data and, the creation of the base map. The data collection process was performed with the help of open data kit (ODK) system (fig - 2). This system helps users to prepare a questionnaire and collect data with the combination of open platform and smartphones. This system enables to collect the geographical location (latitude, longitude and, elevation), and photograph (geocoded) of the case along with other required information.

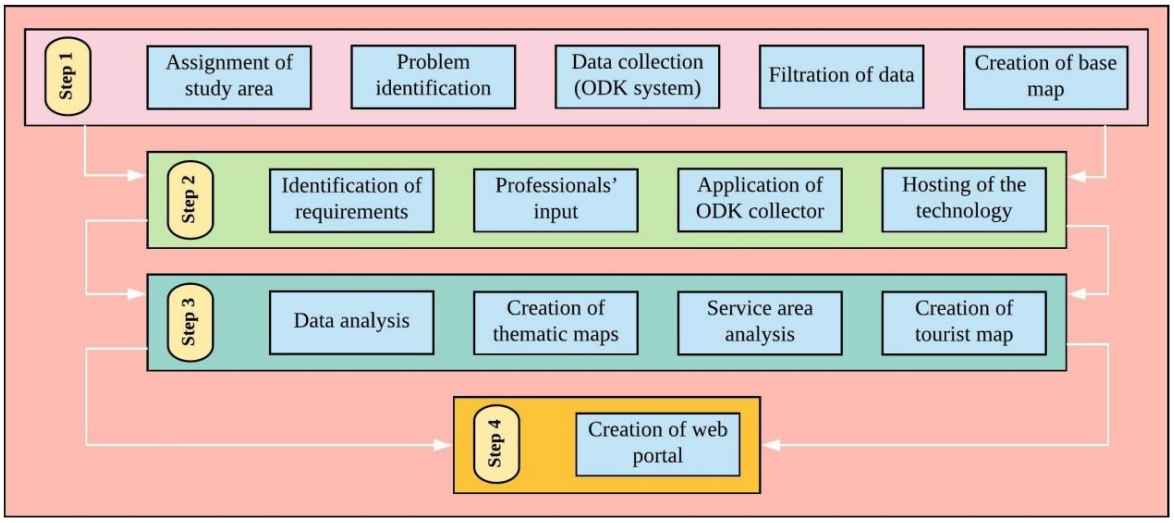

Fig 1: Flow diagram of the research methodology

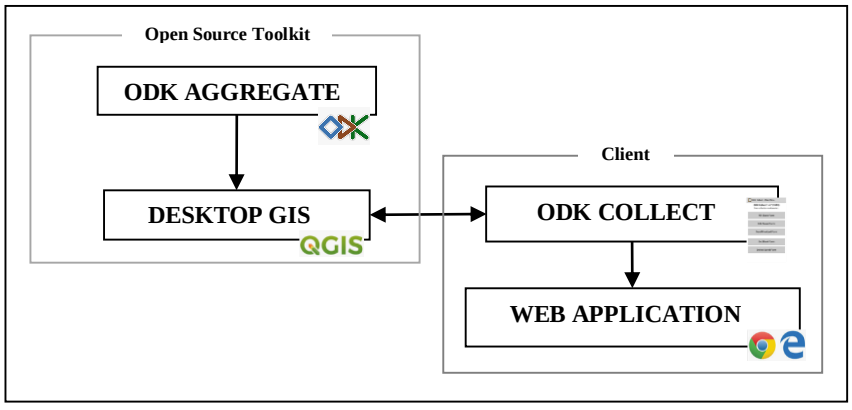

Fig 2: Open source data collection tool

The second part of the methodology includes identification of requirements, professionals' input, and application of ODK collector and hosting of the technology in the system. The third part of the methodology includes the process of data analysis, the creation of thematic maps, network/service area analysis and, the creation of a tourist map. Fourth and the final part of the methodology includes the creation of a web portal which will take filtered data as input. This web portal also will incorporate comments and updates from the users' side. Weather map, open street map, and google map layers will be used for spatial information generation. This web portal will provide the facility for open data integration. This web portal will be used for smart applications in the field of governance, tourism, and citizens & entrepreneurs.

DATA COLLECTION AND ANALYSIS

Data has been collected with the help of ODK system of two essential spots. The first spot lies in Dehradun-Mussoorie road at a distance of 6.5 km from Mussoorie. Total of 18 points of interest (POIs) was collected at this spot. The second spot lies at the mall road of Mussoorie. Total of 31 POIs was collected at this spot. These points were corrected with the help of raster analysis, convex hull analysis, and cluster analysis to remove the biases (wrong data set). For percentage share of the class of data set, please refer Fig. 3. The map was prepared to show the spatial distribution of total collected POIs in QGIS (Fig. 4).

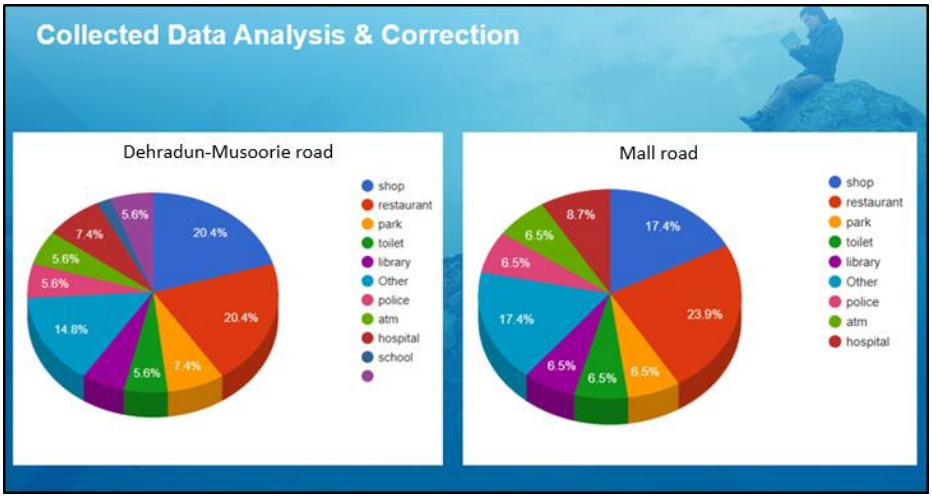

Figure 3: Data before and after correction

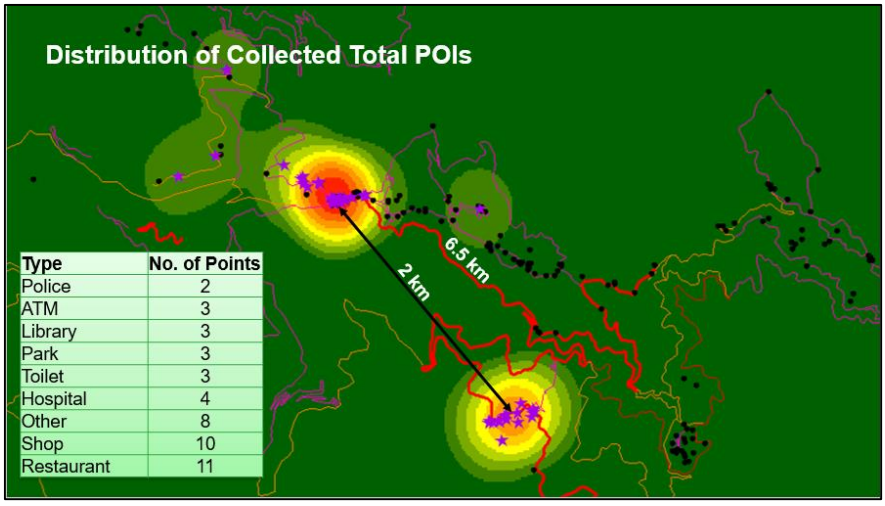

Figure 4: Spatial distribution of total collected POIs

A tourist information map was generated with the help of information gathered in the survey (Fig. 5). In this map not only the available facilities and amenities of the study area has been plotted but at the same time all the essential geocoded photographs have been placed, which can be directly linked to the open online data set, like Google Earth, Google Maps and Google Street and so on. This kind of interlinkage can help us to understand the geographical area, and also the tourist's interest can be coupled many times to visit this place.

Figure 5: Tourist information map

RESULTS AND DISCUSSIONS

The goal of this research paper is to progress understanding of tourist-centric citizen science programs from the sample points collected in the field survey we can assert that the percentage distribution of POIs and tourists' amenities and services are 23.9% restaurants, 17.4% shops, 8.7% hospitals, and medical stores, 6.5% toilets, 6.5% ATMs, 6.5% library, 6.5% parks and 17.4% others. Some problems identified during the survey such as lack of benches and street lights. Most of the benches captured by the vendors.

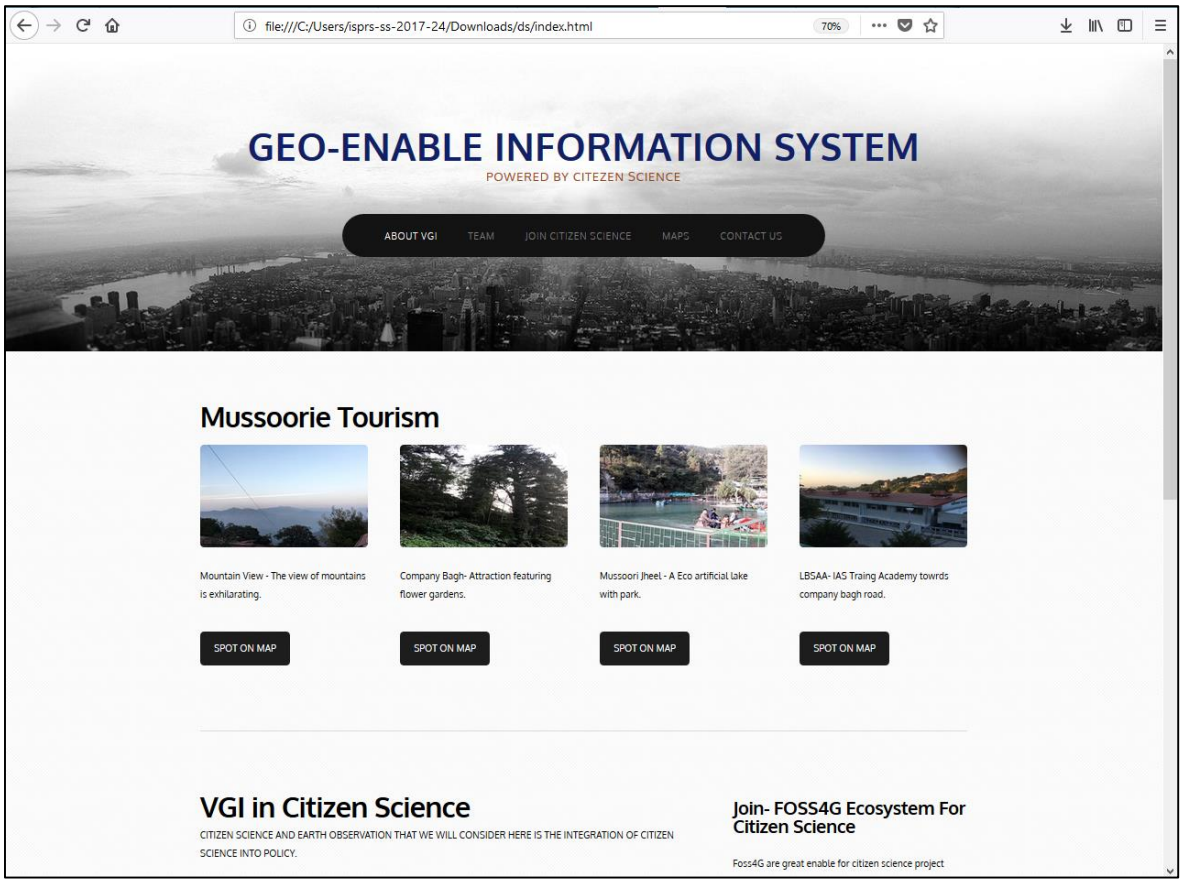

Figure 5: Geo-enable information system

Data points were classified based on distance from the road in QGIS platform. The distribution of the data points are 42 POIs at 50m, 2 POIs at 100m, 2 POIs at 200m and 3 POIs at 500m from the road.

A website was created which integrates the data related to 1) transport, 2) maps/locations, 3) access data/spreadsheets, 4) social media, 5) big data, 6) services, 7) DBMS, 8) sensor networks and, 9) land data. This system will enable any mobile or computational devices, at any geographical location, at any time by any person for data management, planning and analysis, workforce optimization and operational awareness.

This system works on the concept of providing a real-time solution for high-velocity monitoring and alerting. This system also incorporates performance improvements and smart services (Fig. 6).

The significant challenges faced in this study are, (1) Removal of biases from the collected datasets. (2) Accurate location identification (with base maps and GPS points). (3) ISP connectivity problems. (4) Accessibility for data collection.

Therefore, the knowledge of geospatial technology integrated with citizen science (1) helps to make a powerful open map portal for real-time data collection. (2) Data filtering, management, storing, security and privacy is a challenge. (3) These data can be applied to different sectors for getting benefits for all the citizens.

CONCLUSION

Citizen Science creates openings for both volunteers and scientific researchers to expand knowledge, and produce functional scientific data. Investigating the proficiencies of volunteers, characteristics of these data, and measurable outcomes of these programs are a precondition to ensure successful programs for both the volunteers and scientists. Open data connection methods and techniques can use the build an up-to-date geospatial database for providing sustainable solutions to smarter communities. Improvements with mobile technology, comprising software applications and add-ons like hardware sensor device (i.e., Land-Zandstra et al. 2016) consent many different types of data to be collected, elsewhere directly the point location of an animal. Depending on the program, different levels of data gathering may be suitable for diverse volunteers, because mobile technology let this flexibility in citizen science. With the technology developments, citizens move to the smart solutions to solve their day today spatial challenges. The centralized geodatabase would bring advances to Data Management, planning and analysis and make better decisions within a short period using any device from anywhere in the world at any time anyone. In addition to that, data open data collection techniques will be used to train citizens, students, state or private sector staff. Also disaster risk mitigation, tourism, governance, urban infrastructure developments, and environmental studies

ACKNOWLEDGMENT

We thank ISPRS summer school - "Geo-Processing Tools and Technologies in Citizen Science" event conducted by IIRS-ISRO for resources. Especially we heartily thanks to a citizen to Participate in such rigorous work.

REFERENCES

1. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic, Z., Players, F., 2010. Predicting protein structures with a multiplayer online game. Nature 466, 756–760. http://dx.doi.org/10.1038/nature09304.
2. Lintott, C.J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M.J., Nichol, R.C., Szalay, A., Andreescu, D., et al., 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan digital sky survey. Mon. Not. R. Astron. Soc. 389, 1179–1189. http://dx.doi.org/10.1111/j.1365-2966.2008.13689.x.
3. Ellis, R., and C. Waterton. 2004. Environmental citizenship in the making: the participation of volunteer naturalists in UK biological recording and biodiversity policy. Science and Public Policy 31 (2): 95-105.
4. Budruk, M, and DD White. 2008. Connecting Visitors to People and Place: Visitors' Perceptions of Authenticity at Canyon de Chelly National Monument, Arizona. Journal of Heritage Tourism 3 (3): 185–203.
5. Rotman, D., J. Preece, J. Hammock, K. Procita, D. Hansen, C. Parr, D. Lewis, and D.Jacobs. 2012. Dynamic changes in motivation in collaborative citizen-science projects. Page 217. Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW. New York: ACM Press.
6. Land-Zandstra, Anne M, Jeroen L A Devilee, FransSnik, FrankaBuurmeijer, and Jos Mvan den Broek. 2016. Citizen Science on a Smartphone: Participants' Motivations andLearning. Public Understanding of Science 25 (1): 45-60.