

Keys to successful scientific VGI projects

Jens Ingensand¹, Sarah Composto¹, Olivier Ertz², Daniel Rappo², Marion Nappez¹, Timothée Produit¹, Mathias Oberson², Ivo Widmer³, Stéphane Joost³

¹ University of Applied Sciences Western Switzerland, Institute G2C, Yverdon-les-Bains, Switzerland

² University of Applied Sciences Western Switzerland, Institute MEI, Yverdon-les-Bains, Switzerland

³ Ecole Polytechnique Fédérale de Lausanne (EPFL), Geographic Information Systems Lab (LASIG), Lausanne, Switzerland

Abstract

Scientific projects are increasingly using volunteered geographic information (VGI) in order to collect and validate geographic data. This concept relies on the three challenges that A) users can be found and B) be convinced to collaborate and contribute and that scientists finally C) are able to gather high quality data for their projects. In this paper these three challenges are discussed using the experience with three different research projects.

Keywords

Volunteered geographic information (VGI), citizen science, crowdsourcing, crowdcrafting, PPGIS, data quality.

Introduction

Volunteered Geographic Information (VGI) is a recent trend in GIScience that represents an efficient and less expensive way of collecting data. To measure the success of a project both quality (e.g. in terms of precision, completeness, etc.) and quantity of data are important metrics. One key to success are the users [1] [2]: users need to install an application or visit a web-page and thereafter see an interest in contributing to a project. Another important issue is quality [3]. Even if users are willing to contribute to a project, it must be assumed that they are neither experts in the matter they are sharing nor experts in geographic information. In this paper we present findings that we have collected in several VGI projects. The paper is structured as follows: First we briefly present our case studies and research methods, thereafter we will discuss our findings regarding the research questions “how to find users”, “how to motivate and keep users” and “how to get high quality data”. Finally we will present our conclusions and ideas for further research.

VGI case studies and research methods

The findings presented in this paper are based on three projects. Our first case study, called *Urbangene* is a web-based system for the digitization of urban ponds in the greater Geneva area (Figure 2). Citizens were asked to share a pond and to specify which species (e.g. toads) were observed in the pond. The second case study is a project called *Signalez-nous*. The goal of the system is to offer a web-based tool for the digitization of broken streetlights and damaged playgrounds in the city of Yverdon-les-Bains (Figure 1). The third case study is an ongoing project which aims at developing a mobile application called *BioSentiers*. Its main goal is to motivate citizens to share observations and to take actions for biodiversity conservation in the city of Yverdon-les-Bains.

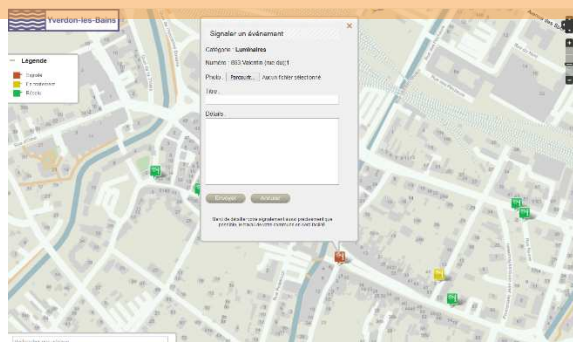


Figure 1: The *Signalez-nous* website

In order to analyze if a system was more successful and why, we used the system's log files and databases (in the case of *Urbangene* and *Signalez-nous*). These two data-sources were able to tell us important measures such as the time and date a user digitized an object, the zoom level at which an object was digitized, the username, the rough position of the user at zip-code level (through IP-to location databases), the users' equipment (e.g. mobile device, type of browser, etc.) and the URL a user visited before using a system.

How to find users

In both projects *Urbangene* and *Signalez-nous* media campaigns were used to get users. Concerning the *Urbangene* project news was spread using newspapers, radio transmissions and social networks, whereas the *Signalez-nous* project was on the state-owned TV, radio transmissions and newspapers. As web-logs showed, each media appearance created usage peaks. Interestingly about a third of the users (identified using the user's distinct ip-addresses) found the *Urbangene* website through social networks which are a less expensive way of finding users.

How to motivate and keep users

In order to make a VGI application successful, users need to continue using an application after discovering it. User motivation can be divided into two main categories [5]: extrinsic motivation is the most common category and depends on reward and punishment. For instance in the case of the *Signalez-nous* project a person contributes to the website because she/he receives a gift at the end of the year (e.g. a solar flashlight). On the other hand intrinsic motivation depends on the user itself and the will to contribute because he/she appreciates using an application or sees a personal interest in using it. Another important point is the contributions' effect. Regarding the *Urbangene* project which is about biodiversity, the effect is less direct and visible as compared to the *Signalez-nous* project. Citizens are more concerned about a damaged streetlight than a pond, because they can feel insecure if streetlights are turned off. On the other hand urban biodiversity has an important, yet less visible and direct effect on a citizen.

An important manner of addressing motivation is recognition. Recognition can be generated through other users (e.g. through comments on a shared objects) or through an automated ranking. In the case of the *BioSentiars* project, a user is offered tools to communicate with other users that allow him to comment on another person's actions. Thereby each person is able to rate other users' contributions.

Another key to keeping users is the optimization of the interface and the minimization of cognitive load [4]. In the case of the *Urbangene* project, the application was kept as simple as possible and uses an adaptive questionnaire that minimizes redundant questions. By decreasing cognitive load, users can concentrate on sharing a pond and specifying which species were observed in the pond. Another way to decreasing cognitive load when displaying spatial content is to use augmented reality. In the case of the *BioSentiers* project, instead of reading and interpreting a map, users see overlaid observations through the camera of a mobile device. Furthermore design should be adapted to the user's needs and skills in terms of technical skills and domain knowledge.

How to get high quality data

There are different ways to check the quality of already digitized data. In the case of *BioSentiers* data is checked not only by specialists through a web application, but also by the community itself. Indeed each user is able to comment, like, confirm or negate each observation. For instance if an observation is negated by many users, it can be considered as false. Moreover data can be cross-verified using reference data; e.g. an observation of a land plant in the middle of a lake will be considered as false.

Furthermore it is essential to help the users to increase data quality; e.g. through tutorials or visible help messages. In the *Urbangene* project, a pop-up is displayed in order to ask the user to zoom in further before digitizing a pond location. In the case of the *Signalez-nous* project a streetlights dataset is displayed as soon as the users zoom in, in order to make the user understand at what level an accurate digitization should be made.

We have found evidence that suggests that authentication increases data quality: in the *Signalez-nous* platform users are required to create an account; in *Urbangene* a user can simply add a username to a digitized object without creating an account. The overall data quality in terms of geometrical precision was significantly better in the *Signalez-nous* platform. Furthermore the administrators of the *Signalez-nous* platform send an email to users who digitizes too much incorrect data.

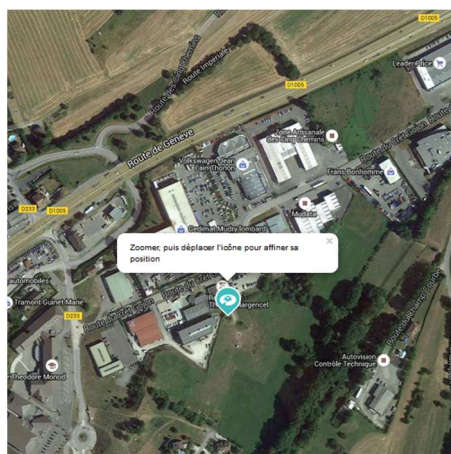


Figure 2: A message indicating to zoom further in order to digitize a more precise location

Conclusions and perspectives

The experience of the three cited project shows that VGI indeed is an interesting possibility for scientific projects to gather high quality data. When designing a mobile application or a website for sharing information, scientists need to imagine who the intended users will be (e.g. using which medium) and at what occasions their attention might be directed to a system. Second it needs to be addressed how the users can be kept motivated to contributing; e.g. through offering a recognition for their work or through manners to communicate with other users. Third the quality of data can be increased if users are offered help in their task to share data (e.g. through tutorials or visible messages). Other manners include specialists validating data, the implementation of automated checks using confirmed reference data or offering tools to all users to validate data.

Further investigations will show what measures are more effective than others and what methods are best used in which context.

References

- [1] J. Preece and B. Schneiderman, «The Reader-to-Leader Framework: Motivating Technology-Mediated Social Participation,» *AIS Transactions on Human-Computer Interaction*, pp. pp. 13-32, 2009.
- [2] M. Haklay, «Citizen Science and Volunteered Geographic Information – overview and typology of participation,» in *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, Berlin, Springer, 2013, pp. 105-122.
- [3] O. Huisman and R. De By, *Principles of Geographic Information Systems*, Enschede: ITC The International Institute for Geo-Information Science and Earth Observation, 2001.
- [4] A. MacEachren, F. Boscoe, D. Haug and L. Pickle, «Geographic Visualization: Designing Manipulable Maps for Exploring Temporally Varying Georeferenced Statistics,» in *Information Visualization*, Research Triangle, Caroline du Nord, 1998.
- [5] D. Berthiaume and N. Rege Colet, *La pédagogie de l'enseignement supérieur : repères théoriques et applications pratiques*, 2014.