The Urbangene Project
Experience from a Crowdsourced Mapping Campaign

Jens Ingensand[1], Marion Nappez[1], Stéphane Joost[2], Ivo Widmer[2], Olivier Ertz[3] and Daniel Rappo[3]
[1] *GIS Lab, Department EC+G, University of Applied Sciences Western Switzerland, Yverdon-les-Bains, Switzerland*
[2] *Laboratory of Geographical Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
[3] *Department Comem+, University of Applied Sciences Western Switzerland, Yverdon-les-Bains, Switzerland*
{*jens.ingensand, marion.nappez, olivier.ertz, daniel.rappo*}*@heig-vd.ch*, {*stephane.joost, ivo.widmer*}*@epfl.ch*

Keywords: Citizen Science, Crowdsourcing, Participatory WebGIS, Social Media, Urban Biodiversity, VGI.

Abstract: Geospatial crowdsourcing applications are emerging systems that enable researchers to collect important information that otherwise would be difficult to obtain. In biodiversity monitoring, crowdsourcing is a promising approach as it benefits from a large group of people with an often underestimated biodiversity and taxonomy knowledge. Despite its huge potential, crowdsourcing approaches are still underrepresented in biodiversity monitoring. We here evaluate a participatory crowdsourcing web mapping platform that was developed to get information about geographic locations and biodiversity characteristics of urban ponds in the Geneva cross-border region. An important fraction of urban ponds is assumed to be located on private grounds, which makes the participatory crowdsourcing approach very valuable. A media campaign was initiated, encouraging citizens participate and to digitize ponds. In this paper we a) evaluate and discuss the impact of the media campaign on the usage behaviour and history of citizens using the crowdsourcing platform and b) assess the quality of the digitized data that has been collected. This study shows that through media campaigns, citizens can be mobilized and motivated to participate in biodiversity crowdsourcing projects. Results indicate that large quantities of users were recruited through social media. However, only a small fraction of about 3% of the mobilized people digitized ponds on the platform. The majority of these users (68%) digitized one pond while 32% digitized two or more ponds. This study shows that it is important for crowdsourcing platforms to be designed and planned in order to facilitate its usage. However, it is crucial for the success of such campaigns to offer something in return to the users and to encourage them to interact among themselves. We suggest that future crowdsourcing biodiversity mapping campaigns should have mobile-optimized interfaces. Mobile devices have the potential to e. g. automatically register coordinates for biodiversity sighting and for uploading respective pictures directly in the field.

1 INTRODUCTION

Spatial crowdsourcing systems are an emerging type of systems. These systems have their roots in the experience from PPGIS (public participatory GIS, Sieber (2006)) and many non-spatial crowdsourcing projects, such as wikipedia (wikipedia.org; Wikimedia Foundation Inc., San Francisco, USA). According to Goodchild and Sui (2011), there are two converging trends that also contribute to the success of spatial crowdsourcing systems: geographic information systems are increasingly social and online social networking sites such as Facebook are increasingly location-based. The success of the OpenStreetMap project (openstreetmap.org; Open Data Commons Open Database License) is an example of a spatial crowdsourcing system where data can potentially be collected and used by everybody. Bartoschek and Kessler (2013) state that most publications related to volunteered geographic information (VGI) focus on this project.

A key challenge is to get people to use a crowdsourcing system or application and to participate in a mapping campaign. This challenge can be divided into two components - on the one hand the targeted citizens need to find the application and on the other hand, citizens must want to use the system once discovered. The incitement to participate in crowdsourced mapping campaigns is often given by existing communities such as, for instance, the global ornithological network *eBird* (Robbins, 2013) or the crisis mapping tool *Ushahidi* which has been successful for

Ingensand J., Nappez M., Joost S., Widmer I., Ertz O. and Rappo D..
The Urbangene Project - Experience from a Crowdsourced Mapping Campaign.
DOI: 10.5220/0005468501780184
In *Proceedings of the 1st International Conference on Geographical Information Systems Theory, Applications and Management* (GISTAM-2015), pages 178-184
ISBN: 978-989-758-099-4
Copyright © 2015 SCITEPRESS (Science and Technology Publications, Lda.)

mapping observations of violence (Liu et al., 2010).

Both projects, Ushahidi and *eBird*, have one thing in common - the participants motivation is already given through the fact that there is a common interest to collect data and to visualize and analyze data that has been collected by others.

However, if there is no existing community available, scientists who want to collect specific data through a crowdsourced mapping campaign face a challenge in getting people to use their system. First of all, potential participants need to be informed of the system's existence e.g. through media campaigns or through social networks. Secondly, the system that is used needs to be optimized for the tasks of sharing and editing spatial data. According to Haklay (2013), it is important that the equipment that volunteers utilize in order to gather information is adapted to their skills and personal knowledge.

Aquatic ecosystems in urban areas are influenced and threatened by human activity, which often causes fragmentation, degradation or loss of environments and reduces connectivity between remaining habitat patches. Ponds and streams are important biodiversity hotspots in urban environments. In order to allow a spatially explicit conservation planning for urban aquatic ecosystems, with the goal to maintain, improve and connect urban aquatic habitats, it is thus important, in a first step, to locate and characterize these vulnerable environments. We evaluated a participatory web mapping platform, which allows for the input and retrieval of information about the location and characteristics of urban ponds. An important fraction of ponds is assumed to be located on private grounds, which makes the participatory crowdsourcing approach very valuable, as such information would be difficult to obtain otherwise.

The platform, called *Urbangene*, allows citizens to digitize ponds and to report specific species that have been identified in or around these ponds. It was developed in the framework of the biodiversity research project Urbangene, giving its name to the platform. The main goal of the Urbangene project is to investigate the impact of urbanization on biodiversity an thereby to allow for more sustainable urban development and the conservation, management and promotion of biodiversity in urban areas.

Urban biodiversity has both a direct and an indirect value for a city's residents; it is for instance important for the physical and psychological health (Aldous, 2007; Fuller and Irvine, 2010). However, urban aquatic biodiversity (such as ponds and its inhabiting species) is less apparent to the citizens than, for example, bird biodiversity. Moreover the mapping of observations of violence Liu et al. (2010) seems to have a more direct gain to the residents of a city. Thus, developers of the Urbangene platform first had to call the attention of the residents to the urban aquatic biodiversity problem through a media campaign. Developers optimized the platforms interface following the most important interface design guidelines (e.g. Shneiderman's eight golden rules of interface design (Shneiderman and Plaisant, 2009)) and developed several mock-ups in order to test and discuss multiple design options. A challenge in establishing the design of the system was to identify the potential end user of the platform.

In the following sections we will focus on the Urbangene biodiversity mapping platform and the media campaign that has been started in order to encourage volunteers to digitize ponds. Thereafter we will analyze the results of the mapping campaign and respond to the research questions:

1. Which medium was more successful in encouraging citizens to use the system? The type of medium can give important indications for further campaigns.
2. Is it possible to characterize the users of the platform? When were they using the system? Where were they located? How many times were they accessing the system? A user profile can help further crowdsourcing campaigns to improve user interfaces and to adjust the systems to the users needs.
3. What is the quality of the digitized data? How accurate are the results? Is there any evidence of fraud? This question is important since it tells if users were willing to collect usable data and if the tool was adapted to the users skills and expectations.

Finally we will discuss the results and give our intentions for further work.

2 THE URBANGENE PROJECT AND PLATFORM

The main aim of the Urbangene project is to investigate the impact of urbanization on biodiversity in the Geneva cross-border area with the goal to allow for more sustainable urban development and the conservation, management and promotion of biodiversity in urban areas. The Geneva cross-border area (Greater Geneva area) includes urban, suburban and rural territories in both Switzerland and France around the city of Geneva.

The main objectives of the Urbangene project are thus the assessment of the current state and the dynamics of biodiversity, the quantification of landscape

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

elements on gene dispersal and functional connectivity of both animal and plant species and, further, the identification of factors that influence gene flow and adaptation of species in urban areas.

In order to collect data, the web-based Urbangene platform had been developed. The focus of this data collection platform was on the identification and localization of species such as toads and newts, and also on their habitats: ponds. Due to the fact that ponds often exist on private property, an important point was to conceive the platform as a crowdsourced data collection system, and thereby to implicate the population of the area. The main function requirement was the digitization of ponds and the specification of species that inhabit the ponds. Users would be able to see the ponds and species digitized by other users and would thus be encouraged to digitize other ponds.

On the client side, the Urbangene platform was developed using HTML5/CSS3 technologies. Mapping functionality was implemented through the JavaScript library Leaflet and enhanced by the jQuery JavaScript library and its numerous extensions. On the server side Apache, PHP and the database system PostgreSQL/PostGIS were used for receiving and storing user input.

Much emphasis has been put on the design of the user interface in order to facilitate the task of digitizing ponds; e.g. an adaptive questionnaire has been implemented that shows questions depending on the answers that the user gave to previous questions or messages that appear in order to guide the users. Figure 1 shows a screenshot of the platform.

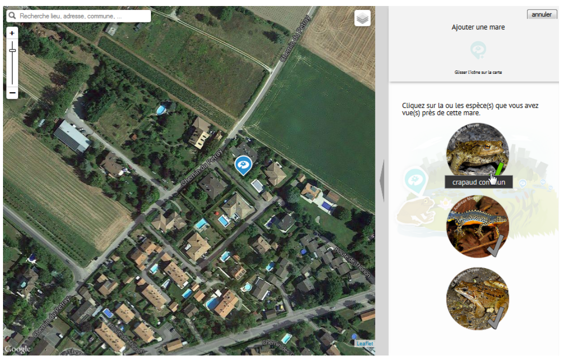

Figure 1: A screenshot of the Urbangene platform.

Once the user gets to the web-based Urbangene platform, a welcome screen is shown. The user then proceeds to the map where the already digitized ponds and the limits of the Greater Geneva area are shown. The system then indicates through a message that a new pond can be added through sliding a pond-icon on the map. Thereby the location of a pond is digitized as a point feature (Figure 1). Once the user has put the icon on the map, the system requires the user to zoom and to check the pond's position. At the same time, the questionnaire is shown and the user can fill in the questions, such as:

- the species identified in the pond (sample images were shown, see Figure 1)
- the distance to a natural space from the user's residence
- how much time a week the user spends in nature

Another important question was whether the user would allow scientists to tissue samples in order to genotype the animals in the case where the pond was located on their property. In addition, the platform allows the user to upload pictures of observed species which could not be clearly identified, and need to be cross-checked by biodiversity experts. Once the questionnaire is completed, the user's feedback is stored in the database and its pond icon is made visible on the map.

3 MEDIA CAMPAIGN

In order to call the citizens attention to the urban biodiversity thematic and to encourage them to participate in the pond survey, a media campaign was started by the university's media and communication center and the creators of the Urbangene project. The media and communication center sent a press release to a large number of newspapers and radio stations throughout Switzerland. A project page and a Facebook page were created, and the link to the platform was published on web pages and forums.

The first articles about the Urbangene platform appeared in four newspapers on the 20th of March 2014; four articles appeared the day after and the national Swiss Radio dedicated a short radio transmission to the project. Overall, news about the Urbangene platform appeared 13 times on Swiss media (see Figure 2).

4 DATA COLLECTION

In order to answer our research questions, we did not only use the platform's database that contained the digitized ponds and the responses to the questionnaire, but also the system's web server log. The Apache log file contains useful information such as the users' ip-addresses, the time and date users accessed the system, the user's browser type and the URL from which the user arrived at the system.

SCIENCE AND TECHNOLOGY PUBLICATIONS

A la chasse aux crapauds du Grand Genève

Le crapaud commun mesure entre 8 et 12 cm, il est trapu et court sur pattes. CHRIS BLASER

Céline Garcin

Une étude cherche à mesurer l'impact de l'urbanisation sur la biodiversité de la région

Après le recensement de la population, le recensement des... crapauds communs. Des chercheurs de l'Ecole polytechnique fédérale de Lausanne (EPFL) s'intéressent à ces petits batraciens dans le cadre d'une étude mesurant l'impact de l'urbanisation sur la biodiversité du Grand Genève. A défaut de pouvoir envoyer un formulaire à chaque crapaud de la région, les scientifiques comptent sur les habitants pour leur signaler*, d'ici au début de l'été, les mares où grouillent ces bêtes à la peau brune couverte de pustules.

Espèce urbaine

«Le crapaud commun nous intéresse parce qu'il s'agit d'une espèce naturellement présente en milieux urbains; elle n'a pas été introduite par l'homme, explique Stéphane Joost, chercheur à l'EPFL en charge du projet. Notre but est de comprendre ce qu'il se passe génétiquement chez ces animaux lorsque leur population est divisée en petits groupes en raison de l'apparition dans le paysage de nouvelles routes ou bâtiments.»

Pour déceler ces changements, les scientifiques étudieront l'ADN de plus de 200 crapauds de la région. A l'aide de marqueurs génétiques, ils observeront la diversité des populations analysées. «Plus la diversité génétique est élevée, plus l'animal a de chances de résister à différents types de pressions environnementales», résume Stéphane Joost.

Mais pour conserver un bagage génétique important, les différentes populations de batraciens doivent pouvoir se rencontrer. «La connexion avec des populations situées à l'extérieur de la zone urbaine dense est nécessaire pour la

Figure 2: Extract from a newspaper article encouraging readers to digitize ponds on the Urbangene biodiversity mapping platform. (Tribune de Geneve, April 2, 2014).

The user's ip-address can give information about the Internet provider of the user and thus if users were accessing the system using a wired Internet connection or through a mobile 3G or 4G network. Moreover ip2location databases (IP2Location.com, Pulau Pinang, Malaysia) can give an approximate location (e.g. at zip-code level) of the ip-address. The user's browser type allows to analyze if the user was using a mobile device such as a tablet computer or a smartphone; in some cases it also gives information about the operating system installed on the device (e.g. type and version). The web page the user was visiting before clicking on the platforms URL (the referrer) gives an indication how the user found the system; e.g. through a search-engine, through a newspaper's web page or through social media.

We parsed the Apache log file into a database and aligned its data to the data that had been stored in the database through the platform. This new database was this study's main source of information.

5 RESULTS

5.1 Data Collected

The data collected on the Urbangene platform stretches over a period of five months; from the 20th of March until end of August 2014. The main data for this analysis can be characterized as follows:

- After cleaning the log file for robots, test users and other irregular data, we were able to identify 933 distinct IP addresses that used the system during the given period of time.
- 53 ponds were digitized within the limits of the Greater Geneva area.
- 34 distinct users digitized the 53 ponds.
- Most digitized ponds were located in and around the city of Geneva (within the borders of Switzerland). Only six ponds were digitized in France. Figure 3 shows the locations of the ponds.

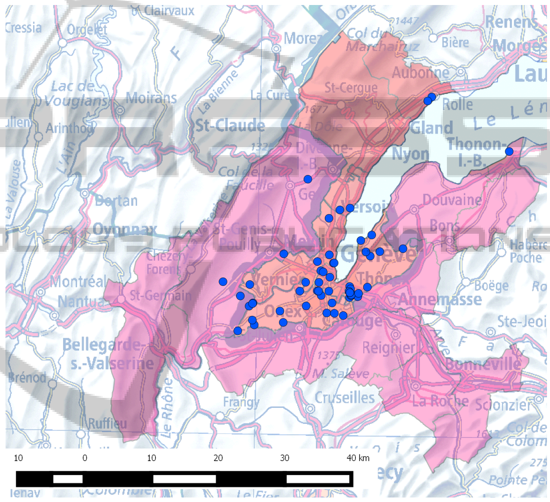

Figure 3: Map showing the results. Blue dots: ponds that have been located and described by citizens using the Urbangene biodiversity mapping platform. Pink area: the French part of the Greater Geneva area, Orange area: the Swiss part of the Greater Geneva area.

5.2 Effectiveness of the Media Campaign

In order to measure the effectiveness of the media campaign, we analyzed the referrers in the log file (i.e. the page the user was visiting before connecting to the platform). Figure 4 shows the results of this analysis. Twenty-five percent of the users reached the platform through the project's homepage, 12% through a media's (newspaper or radio station homepage) and 11% through a search engine. These numbers added together suggest that almost half of the users (48%) found the system through either newspaper or radio. 18% of the users found the URL of the system through other homepages (e.g forums, etc) and 34% of the users came from social networks such as Facebook and Twitter. The high number of social media users was surprising since the Facebook page

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

only had about 30 followers at the time of writing this paragraph.

The fact that only six ponds were digitized in France could be an indication that the media campaign did not sufficiently reach many citizens living in France - however, the part of the Greater Geneva region that is situated in France is partly mountainous and much less populated than Geneva, the second most populated city in Switzerland.

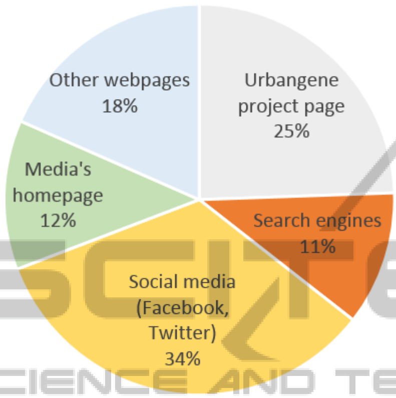

Figure 4: How the users found the platform, i.e. last Internet pages users visited before connecting to the Urbangene biodiversity mapping platform.

We also analyzed on which days the users used the system and when the users digitized ponds. Figure 5 shows clear peaks at around the 20th and 21st of March. Interestingly most of the ponds were digitized a few days after the first articles appeared. These results lead to the argumentation that the media campaign quickly motivated many users to take a look at the system; however, the interest faded after the first week. The observation that users still digitized ponds some days after the first peaks appeared, might be interpreted that these users either were looking at the newly indicated pond locations for the occurrence of different species that were demanded in the questionnaire or went home after discovering the platform and then digitized a pond.

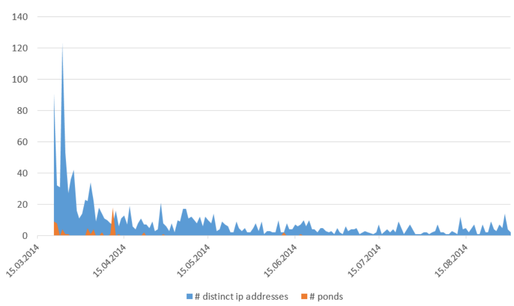

Figure 5: Number of different IP-addresses that accessed the Urbangene biodiversity mapping system and number of digitized ponds from March to August 2014.

5.3 Characterization of the Users of the System

Most users (59%) who accessed the system (we are referring to the 933 distinct IP addresses) were using a desktop or laptop computer, 8% a mobile system such as a smartphone or a tablet computer and 33 % were using a non-identified system.

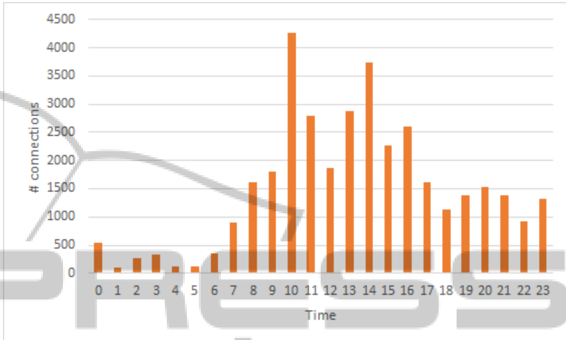

Figure 6: Number of platform requests according to the time of day.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

Most users used the platform during office hours with peaks at 10 a.m and 2 p.m. (see Figure 6); an analysis of the Internet provider of the users showed many different companies. A possible interpretation of these numbers is that people used the system at their workplace, after drinking a coffee and reading the newspaper at 9 or 10 a.m. or after having lunch. This hypotheses can be strengthened by an analysis of the day of week people accessed the system - the data showed that on average only 9% of the traffic generated on the site occurred on a Saturday or Sunday, while 15% - 18% occurred on a working day. The vast majority of users (68%) digitized one pond; few users digitized two or three ponds and only one user digitized six ponds. (see Figure 7).

The ip2location database that we used for identifying the position of the users gave an interesting result for the users who digitized ponds: nine out of 34 users (26%) were located in the canton of Geneva, three in France and seven in other Swiss cantons. 15 users (44%) were using a mobile broadband connection and it was thus not possible to identify a more precise location than Switzerland. A presumably high percentage of mobile device users was also confirmed when we analyzed the browser type of the users who digitized ponds - at least seven users (21%) were using a tablet computer.

More than half of the users who digitized a pond (18 out of 33) answered that a natural space can be found around their house; 5 out of 33 replied that nature is close to where they live (some hundred me-

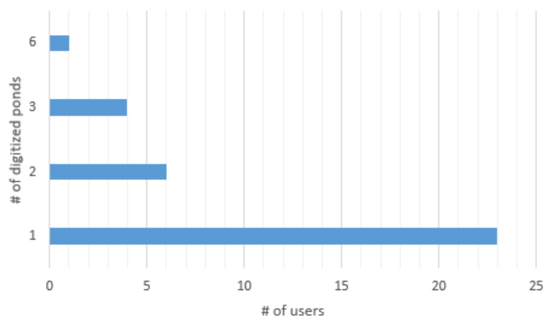

Figure 7: Number of digitized ponds per user.

ters) and seven people did not answer the questions at all. 10 people were spending at least one hour in a natural space a day, while 13 were spending at least some hours a week. The most common species that had been identified was the common toad (bufo bufo) (found in 33 ponds). The common frog (rana temporaria) had been identified in 17 ponds and the alpine newt (ichthyosaura alpestris) in 16 ponds.

5.4 Quality of the Digitized Ponds

The locations of the 53 digitized ponds were compared to other data sources in order to verify its accuracy. For the Swiss part of the Greater Geneva area we used a pond database that had been established by the Geneva Institute of Technology, Architecture and Landscape (hepia). Moreover, we used OpenStreetMap data and Google Maps imagery in order to identify ponds that were not contained in the reference database from hepia.

These three reference data-sources confirmed the location of 26 ponds (49%), 19 locations were not confirmed and thus possibly errors (36%) and 8 locations (15%) were either not far from a confirmed pond (and thus perhaps only an imprecision), not visible due to presumably covering trees, or objects that are other water objects (e.g. fountains or swimming pools)

6 DISCUSSION

We here analyzed the impact of a media campaign on the usage behavior and history of citizens using the Urbangene mapping platform. Although the area of investigation was located in Switzerland and France, the media campaign focused mainly on Switzerland. This fact possibly explains why only few ponds were digitized in France. The impact of the media campaign could be seen in the fact that several hundred people visited the mapping platform.

However, in relation to the effort of the media campaign, the fact that only 3% of the users digitized 53 ponds might seem poor, moreover the traffic on the site quickly declined after only a few days. Having this in mind and along with the fact that most users only digitized one single pond leads to the conclusion that most users did not come back to check how the project was evolving.

There are many motivational reasons for one user to become a contributor (Shneiderman and Preece, 2009), yet one reason explanation for our observation might be found in Haklay (2013) who states that public participation can be categorized into four levels ranging from crowdsourcing (i.e. citizens that are used as sensors) on the lowest level, distributed intelligence, participatory science and extreme citizen science (e.g. citizen participating in the problem definition) on the highest level. In crowdsourcing citizens are thereby used as data collectors and the sum of the digitized data are not of direct interest to the users.

On the other hand, many successful crowdsourcing systems such as OpenStreetMap offer something in return for the work of the users: this is the possibility for one individual to reuse the data added by many users. Another key for successful crowdsourcing biodiversity campaigns appears to be related to the level of networking and organization of communities that are interested in specific species. These communities are often already well-organized and established and in such cases crowdsourcing tools help to gather data that the whole community is interested in.

The fact that many users found the platform through social media is encouraging. The effort to arrange Facebook pages or to tweet URLs is less than starting a campaign including print media and radio stations. Furthermore it is less demanding for the user to click on a URL that is communicated through social media (e.g. a Facebook page) than retyping an address that is printed in an article or communicated during a radio transmission.

7 CONCLUSIONS AND FURTHER WORK

This study indicates that through media campaigns, many citizens can be mobilized and motivated to participate in biodiversity crowdsourcing projects.

Even though it is important for crowdsourcing platforms to be designed and planned in order to facilitate its usage, what is more crucial for the success of such campaigns is to offer something in return to the users. This would encourage users to interact more among themselves in checking the data added

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

by other users and possibly in adding new data to the platform in the future.

Another important lesson learned is that social media is able to recruit large quantities of users with relatively little effort.

We suggest that future crowdsourced mapping campaigns should have mobile-optimized interfaces. Mobile devices such as smartphones or tablets are promising tools for biodiversity monitoring during outdoor activities. Applications for such systems have the potential to for example automatically register coordinates for biodiversity sighting and for uploading respective pictures directly in the field. The usage of mobile devices is increasing and users could be recruited to spend some time using a crowdsourced system during their journey to work or other activities during which they could use their mobile devices. Another advantage is that most social media platforms such as Facebook or Twitter are available on mobile devices. Thereby users could be recruited through mobile social media applications and then directed to crowdsourced mapping platforms without the need to switch device.

ACKNOWLEDGEMENTS

The authors would like to thank all the participants who used the Urbangene platform and the Geneva Institute of Technology, Architecture and Landscape (hepia) for providing the pond-database.

REFERENCES

Aldous, D. (2007). Social, environmental, economic, and health benefits of green spaces. In *Proceedings of the International Symposium on Horticultural Plants in Urban and Peri-Urban Life*, pages 171–85.

Bartoschek and Kessler (2013). *Crowdsourcing Geographic Knowledge*, chapter VGI in Education - from K-12 to Graduate Studies, pages 341–360. Springer.

Fuller, R. and Irvine, K. (2010). Interactions between people and nature in urban environments. *Urban Ecology; Gaston, K .J., Ed., Cambridge University Press: Cambridge, UK,*, pages 134–171.

Goodchild and Sui (2011). The convergence of GIS and social media: challenges for GISscience. *International Journal of Geographical Information Science (IJGIS)*, 25(11):1737–1748.

Haklay, M. (2013). Citizen science and volunteered geographic information: overview and typology of participation. In Sui, D., Elwood, S., and Goodchild, M., editors, *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, pages 105–122. Springer.

Liu, S., Iacucci, A., and Meier, P. (2010). Ushahidi Haiti and Chile: next generation crisis mapping. *ACSM Bulletin*, 246.

Robbins, J. (2013). Crowdsourcing, for the birds. *The New York Times*.

Shneiderman, B. and Plaisant, C. (2009). *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Pearson Addison Wesley.

Shneiderman, B. and Preece, J. (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 1(1):13–32.

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers*, 96 (3):491–507.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS