Chapter 5
Crowdsourced Data

Karel Charvát and Michal Kepka

Abstract Crowdsourcing together with Volunteered Geographic Information (VGI) are currently part of a broader concept – Citizens Science. The methods provide information on existing geospatial data or is a part of data collection from geolocated devices. They enable opening parts of scientific work to the general public. DataBio Crowdsourcing Solution is a combination of the SensLog server platform and HSLayers web and mobile applications. SensLog is a server system for managing sensor data, volunteered geographic information and other geospatial data. Web and mobile applications are used to collect and visualize SensLog data. SensLog data model builds on the Observations & Measurements conceptual model from ISO 19156 and includes additional sections, e.g., for user authentication or volunteered geographic information (VGI) collection. It uses PostgreSQL database with PostGIS for data storage and several API endpoints.

5.1 Introduction

Crowdsourcing is a sourcing model in which individuals or organizations obtain goods and services, including ideas, voting, micro-tasks and finances, from a large, relatively open and often rapidly evolving group of participants. Crowdsourcing can be used as a research method (Citizens Science [1]), as the involvement of the public in scientific research [2].

In the area of collection of spatial information or Earth Observation, we are often using the term Citizens' Observatories [3]. This term is usually understood as methods of community-based monitoring using novel Earth Observation applications and sensors embedded in portable or mobile personal devices [4–6].

K. Charvát (✉)
Lesprojekt-Služby, Ltd, Záryby, Czech Republic
e-mail: Charvat@lesproject.cz

M. Kepka
University of West Bohemia, Pilsen, Czech Republic

C. Södergård et al. (eds.), *Big Data in Bioeconomy*,
https://doi.org/10.1007/978-3-030-71069-9_5

Another term, which is often used in this context, is Volunteered Geographic Information (VGI) [7], which is the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals [8]. Some examples of this phenomenon are WikiMapia, OpenStreetMap, and Google Map Maker. VGI can also be seen as an extension of critical and participatory approaches to geographic information systems and as a specific topic within online or web reliability. These sites provide general base map information and allow users to create their own content by marking locations where various events occurred or certain features exist. In voluntary data collection, an important part is how data is processed. An example is Neogeography (New Age Geography) focused on combining geotagged data (e.g. Keyhole Markup Language—KML) [9] with a map interface for contextualised exploration.

In this chapter we will describe two concepts developed in the DataBio project. Firstly, a solution based on SensLog [10] and a profile for VGI. The second concept is a Map composition and sharing Maps as objects among users.

5.2 SensLog VGI Profile

SensLog is a web-based solution for receiving, storing and publishing sensor data of different kinds. As VGI can be collected as sensor data, SensLog can provide a suitable operational solution. The SensLog data model was extended with new tables with emphasis on the variability of VGI. Only a few mandatory attributes characterize an VGI observation, but it can be enriched with a lot of additional attributes. The data type of an additional attribute is only limited to those that can be stored as a value in JSON format. A VGI observation can include a list of multimedia files that are also stored in a data model. The data model of a VGI module is shown on Fig. 5.1.

Added tables are following:

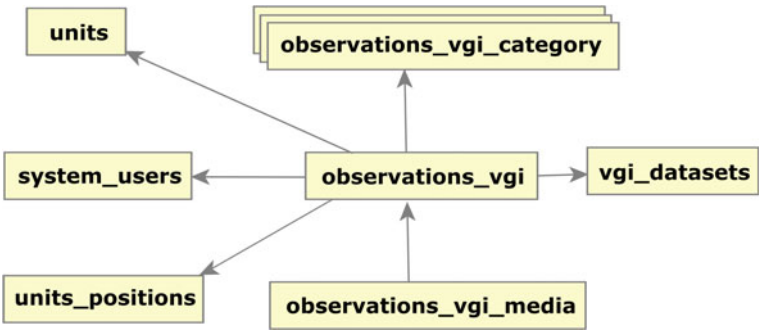

Fig. 5.1 SensLog VGI module data model

- *observations_vgi*—a main table storing VGI observation with all connected attributes
- *observations_vgi_media*—a table storing multimedia files connected to VGI observations
- *observations_vgi_category*—a table storing coded list values of categories of VGI, uses partitioning mechanism to sort categories
- *vgi_datasets*—a table storing user-defined datasets of VGI observations with metadata

This data model supports measurements and observations by users with portable devices. It is a typical way of collecting spatially referenced thematic data in the domain of the Earth Observation (Fig. 5.2).

To publish data according to Open Linked Data best practices and with a self describing data structure, we are using Virtuoso graph model engine [11]. The data are stored in so called *quads*, which consist of a graph name describing a dataset and triple mapping an attribute value (object) to Point of Interest (POI) (subject). The quad uses a property which, in the best case, is well defined in some public ontology making it easier to integrate our data into other systems [12].

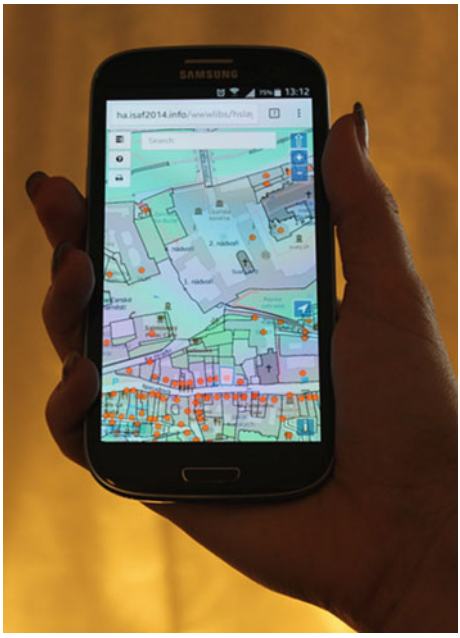

Fig. 5.2 Citizens observatories mobile Apps

5.3 Maps as Citizens Science Objects

Maps are interesting not only as visualizations of agriculture data—but also as shareable, fascinating and valuable agriculture objects in themselves. In the past, a map used to be an expensive rolled up scroll of calf skin that was drawn by a skilled artist from the manuscripts of daring sea-farers in the great age of discovery. Later, maps were produced by less picturesque but more efficient means until the advent of the Geographical Information Systems (GIS) age, when a lot of people suddenly could make professionally LOOKING maps. Nowadays, a map is not a "flat image", but a complex layered object that references data sources 'scattered' across a decentralized, democratic and, at times, volatile Internet.

Our needs are many and very different, but so are our skill sets. Thus, offering everyone sophisticated GIS tools capable of letting the users make their own maps is not the best way. It is often simpler, better and more effective to simply give them a "map".

Currently, hundreds of services offer spatial information through real-time interactive protocols such as Web Map Service (WMS) and Web Feature Service (WFS) etc. Soon, if EU member states and signatories of INSPIRE do as they are legally obliged, this number will be thousands, even ten thousands.

As a map is a composite object referring to a lot of live data sources around the internet, it requires a "Map Composition" standard that describes the map elements and how they should be combined to fit together neatly.

An early effort by the OGC was the Web Map Context specification that has not evolved since 2005. This slightly "heavy-weight" XML-based standard is limited in scope and has not kept up with the developments in standards and technology in the years that have passed since its creation. In DataBio we worked with defining a simple, lightweight specification for Map Compositions using HTML5 and bandwidth friendly JavaScript Object Notation (JSON) as a carrier of information.

The current specification of the JSON Map Composition is available on the GitHub Wiki of HSLayers NG [13].

References

1. Irwin, A. (1995). Citizen science: A study of people, expertise and sustainable development. Routledge. ISBN 9780415130103.
2. Gura, T. (2013). Citizen science: Amateur experts. *Nature, 496*(7444), 259–261. https://doi.org/10.1038/nj7444-259a.
3. Have you heard about the concept of Citizens' Observatories?, Published on 21/12/2016. Available at: https://ec.europa.eu/easme/en/news/have-you-heard-about-concept-citizens-observatories.
4. Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., & Crowston, K. (2012). The future of citizen science: Emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment, 10*(6), 298–304.
5. Robinson, J. A., Kocman, D., Horvat, M., & Bartonova, A. (2018). End-user feedback on a low-cost portable air quality sensor system—are we there yet? *Sensors, 18*(11), 3768.

6. Charvat, K., Bye, B. L., Mildorf, T., Berre, A. J., & Jedlicka, K. (2018). Open data, VGI and citizen observatories INSPIRE hackathon. *International Journal of Spatial Data Infrastructures Research, 13,* 109–130.
7. Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*(4), 211–221.
8. Elwood, S. (2008). Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal, 72*(3&4), 173–183.
9. Graham, M. (2010). Neogeography and the palimpsests of place. *Tijdschrift voor Economische en Sociale Geografie, 101*(4), 422–436.
10. SensLog. http://www.senslog.org/.
11. Virtuoso Universal Server. https://virtuoso.openlinksw.com/.
12. Kepka, M. et al. (2017). The senslog platform—a solution for sensors and citizen observatories. In J. Hřebíček, R. Denzer, G. Schimak, T. Pitner (Eds.) *Environmental software systems. Computer science for environmental protection.* ISESS 2017. IFIP Advances in Information and Communication Technology, Vol. 507. Springer, https://doi.org/10.1007/978-3-319-89935-0_31.
13. Patras DataBio Hackathon TEAM 4: Sharing maps as intelligent objects. https://www.plan4all.eu/2019/06/team-4-sharing-maps-as-intelligent-objects/.