Haklay, M., 2017. Volunteered Geographic Information and Citizen Science, in Kitchin, R., Lauriault, T.P., and Wilson, M. (eds.), *Understanding Spatial Media*. London: Sage, pp. 127-135

Chapter 13 - Volunteered Geographic Information and Citizen Science

Muki Haklay (University College London)

13.1 Introduction

In this chapter, we explore the related areas of Volunteered Geographic Information (VGI) and citizen science. For our purpose here, we will define VGI as digital geographical information that is generated and shared by individuals. VGI can be viewed as the part of user-generated content which has become a major element of Web media over the past two decades. Within VGI, geographical information is an integral part of the digital media object – for example coordinates as an integral part of the Exchangeable image file format (Exif) element of a picture taken with a digital camera (Goodchild, 2007). Citizen science, on the other hand, is defined by the Oxford English Dictionary as 'scientific work undertaken by members of the general public, often in collaboration with or under the direction of professional scientists and scientific institutions' (2014). Citizen science can also be considered as a type of user-generated content, whereas this content refers to scientific facts, observations or analysis.

The foundational concept, that of user-generated content, started early on in the World Wide Web (Web), when systems such as GeoCities (launched in 1994) allowed people to link to the system over the Internet and create their own websites, even though they possessed relatively limited technical skills (Brown, 2001). With further technological and interaction design advances, it became possible to create content with even less technical knowledge through weblogs (Blogs), images, audio (podcasts) and video sharing websites. Around 2005, due to technical, societal and organisational changes, the process of capturing geographic information using a range of affordable devices – from Global Positioning System (GPS) receivers through to cameras and phones – became increasingly available to a wider group of people (see detailed analysis of the underlying trends at Haklay et al., 2008). Of these changes, it is worth noticing one concept in particular within a societal and organisational realm: crowdsourcing. The term 'crowdsourcing' was coined by Howe (2006) to describe a process in which a large group of people is asked to perform business functions that are either difficult to automate or expensive to implement. Fundamentally, crowdsourcing allows an organisation (be it a company or a scientific research institute) to ask a large group of unremunerated or marginally remunerated people to carry out piece work tasks for which the organisation is the prime beneficiary.

Of particular importance, is the labelling of a purposeful activity that people can partake in and the use of terms such as 'volunteer', 'citizen', 'user' and 'crowd' to describe the participants. While 'volunteer' and 'citizen' are clearly loaded with meaning, 'crowd' and 'user' might seem neutral or simply descriptive. Yet, crowdsourcing has been well criticised as an exploitative practice that reduces humans to automatons or machine parts (Silverman, 2014) and therefore 'crowd' is used to treat the contributors to the activity as an anonymous,

faceless (and potentially expendable) group, while the term 'user' in digital technology has been criticised by Brenda Laurel (2001), who observed that 'user'

> *'… implies an unbalanced power relationship – the experts make things; everybody else is just a user. People don't like to think of themselves as users. We like to see ourselves as creative, energetic beings who put out as much as we take in'*

and she goes on to suggest the term:

> *'Partner – this person has agreed to work on something together with you. The idea of being in partnership with the people purchasing your products or on your site is not only emotionally attractive; it is quite literally true'* (p. 49-50).

As for 'volunteer' in VGI, this has received special attention by Sieber and Haklay (2015), who argue that the assumption of free will volunteering, without any wish for personal gain, is not reflected in practices such as crowdsourcing where there is no explicit volunteering for a higher cause and, conversely, instead of seeing volunteering as a reason to increase the trust in the participant, it is a source of concern about their motivations. As can be expected, the 'citizen' in citizen science has also raised a lot of questions as demonstrated in Mueller et al. (2012) who argue that the use of the term 'citizen' requires the linking of public participation in science to a strong concept of democratisation and citizenship, especially when citizen science projects are linked to education. Similar sentiments are echoed in the responses to their paper by Cooper (2012) and Calabrese Barton (2012).

These are merely a few examples of a much wider literature that critiques and questions the use of these loaded terms to describe large-scale activities that have emerged in the past decade. Arguably, they are the result of the underlying tensions that are at the basis of VGI and citizen science, as either altruistic, collaborative efforts towards a common goal and a greater good on the one hand and extracting free labour, in an exploitative way, where the benefits justifiably accrue to the entrepreneurs who have set up the system or have the knowledge and skills to exploit the resulting information on the other. The reality is, as expected, somewhere in between, depending on the nature of the project and its dynamics.

Beyond the discussions about the terminology and its meaning, we should notice the scale and reach of these activities which engage millions of people across the world through data collection, information sharing and the analysis of geographic information. In this chapter we look at the activities that fall under the banners of VGI and citizen science, and especially the intersection between them. Throughout the chapter, we will explore the intention of volunteers in their act of participation and the issue of power between the contributor and the technical and social systems that facilitate the contribution, and will emphasise the values that are embedded into the practices of participation and information sharing.

13.2 Volunteered Geographic Information and Citizen Science

This section examines the several core characteristics of the fields of volunteered geographic information (VGI) and citizen science and provides some cases to demonstrate a range of activities.

The framework used as a basis for analysis is from Craglia, Ostermann and Spinsanti (2012) in which they suggest differentiating between *volunteering* and *geographical* content. They further suggest differentiating between *implicit* and *explicit* contributions. *Explicit volunteering* is when people are knowingly volunteering effort to a project, while *implicit volunteering* is when information is shared openly, but without people knowing how their contribution will be eventually used. For example, carrying out bird observations and reporting to a shared database is considered to be active volunteering, while the reuse of all the georeferenced images of parks that are shared on Flickr to assess the level of interest in the parks or the cleanliness of these parks is implicit volunteering, since the images were shared without this purpose in mind.

Another helpful distinction can be made whether or not the participant needs to *actively and knowingly* contribute information (e.g. use an app such as WideNoise to measure the level of noise – see Becker et al. 2013) or *passively* share (e.g. use a phone to sense the signal from different telephone masts and share this information on OpenSignal[1]).

Volunteered Geographic Information (VGI)

Within VGI, there are activities that clearly fall outside the realm of citizen science – for example, when people use their phone to 'check in' to a bar or provide a restaurant review in apps such as Yelp.[2] This is an example of explicitly volunteering, and an active and explicit geographic contribution. VGI also includes contributions to Wikipedia[3] that contain place names (e.g. an article about a historical figure mentioning a place that they travelled to) but are not explicit geographic contributions, since the aim of the article is not geographical, although it is explicitly contributed.

Some VGI is very similar to citizen science in that it is concerned with recording geographical facts and observations. An application such as StreetBump[4] runs on a participant's phone, who uses the sensors to detect bumps in the road while they are driving their car. Here, there is an explicit sensing of car movement associated with the geographic location from GPS. This is explicitly volunteered, passive and an explicit geographic contribution. OpenStreetMap[5] is another VGI example that parallels citizen science as it is concerned with recording facts about the world and measuring them accurately.

Citizen Science

[1] http://www.opensignal.com/
[2] http://www.yelp.com/
[3] http://www.wikipedia.org/
[4] http://www.streetbump.org/
[5] http://www.openstreetmap.org/

The following willbriefly look at the types of activities that are included in citizen science (for a more in depth examination see Haklay 2013, 2015). Six types of citizen science are discussed here: 1) passive sensing, 2) volunteer computing, 3) volunteer thinking, 4) environmental and ecological observations, 5) participatory sensing and 6) community science.

1) **Passive sensing:** relies on participants in the project providing a resource that they own (e.g. their phone) for automatic sensing. The information that is collected through the sensors is then used by scientists for analysis (ex. StreetBump).
2) **Volunteer computing:** is a method in which participants share their unused computing resources, on their personal computer, tablet or smartphone, and allow scientists to run complex computer models during the times when the device is not in use; (ex. when people use the sensor in their laptop to augment seismographic networks in the Quake-Catcher[6] project).
3) **Volunteer thinking:** uses what Clay Shirky (2011) termed 'Cognitive Surplus', which is the cognitive ability of people not used in passive leisure activities such as watching TV. In this type of project, the participants contribute their ability to recognise patterns or analyse information that will then be used in a scientific project. (ex. GeoTag-X[7] recruits volunteers to help with classification of images as part of humanitarian efforts).
4) **Environmental and ecological observation:** focuses on monitoring environmental pollution or observations of flora and fauna, through activities (ex. bio-blitz in which a group of volunteers study a site thoroughly, using their phones to record and share observations).
5) **Participatory sensing:** is similar to the previous type of observation, but gives the participant more roles and control over the process. While many environmental and ecological observations follow data collection protocols that were designed by scientists, in participatory sensing the process is more distributed and emphasises the active involvement of the participants in setting what will be collected and analysed (ex.WideNoise)
6) **Community/Civic science**, also known as bottom-up science, is initiated and driven by a group of participants who identify a problem that is a concern for them and address it using scientific methods and tools. Within this type of activity, the problem, data collection and analysis are often carried out by community members or in collaboration with scientists or established laboratories (ex. community led air quality monitoring where the check if a local factory is polluting the local environment).

When examining the overlap between citizen science activities and geography (Figure 13.1), the activities of passive sensing, participatory sensing, environmental and ecological observations, and civic/community science inherently happen in a geographic location, and are part of VGI can be seen. When passive sensing focuses on health issues it might not use geographical information, and therefore it is depicted as potentially non-VGI. The case is

[6] http://qcn.stanford.edu/
[7] http://geotagx.org/

more complex with volunteer computing and volunteer thinking, where projects do not necessarily deal with geographic information and can be about analysing neurons or looking at images of galaxies. Here, only when the issue is explicitly geographic – as in classifying images from a camera trap in the Serengeti – is the end result VGI.

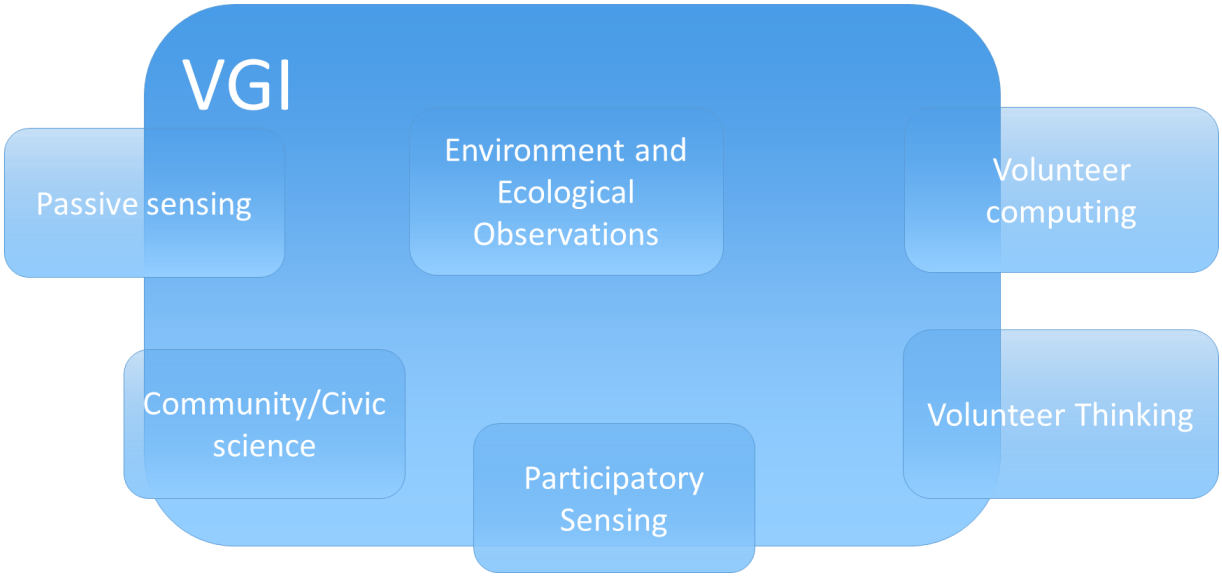

Figure 13.1- Conceptual overlap between VGI and Citizen Science

13.3 Stakeholders and roles in VGI and Citizen Science

Within the areas of VGI and citizen science, different actors take part in creating the systems that facilitates data collection and sharing. First, in VGI there are few projects established without explicit profit motives, and these are run by volunteer-based organisations. OpenStreetMap, for example, is run by the OpenStreetMap Foundation, and it is the exception rather than the rule. Because of the complexity of setting up a system and running it efficiently, many VGI activities considered to be a form of crowdsourcing and are run by companies who usually ensure that they preserve some or all intellectual property rights over the contributions. For example, although the images on Flickr are owned by the people who upload them to the service, Yahoo! holds some rights over what can be done and control access. The imbalance of power and access to information becomes clear as here each contributor is treated as an individual, and their contribution is miniscule compared to the totality of the information that the system owner has amassed through the effort of the participants. The system that facilitates crowdsourcing has costs associated with that provision, and for most participants, the contribution is comparable to the benefits that they extract from the system. For example, when a participant passively submits, location information from a satellite navigation device and the service provider delivers warnings of traffic delays on the basis of these reports, there is a balance between the minimal cost of sending the information and the service that the driver receives. Here, it would be correct to question whether or not the service provider should release all the information voluntarily provided by drivers at their own expense. And what of participants who invest significant amounts of time contributing to the project, is it also appropriate to ask about the obligation of the service provider to such a contributor? The techno-libertarian answer would be that the volunteer does that of their own free will and the service provider doesn't have any responsibility to her/him. These are but a few moral dilemmas.

Citizen science, is a more interesting and varied scenario in terms of the organisations that coordinate citizen science activities (see Haklay, 2015). To date, there has been little commercial interest in citizen science activities and science is perceived as an activity that aims to improve the sum of human knowledge, even if there are specific individuals who benefit from it from this activity more than others. Historically, citizen science is an area that has involved many charities and non-governmental organisations, and one of the most celebrated examples of citizen science – the Christmas Bird Count (see Goodchild 2007) – is run by the National Audubon Society, a US not-for-profit dedicated to conservation. The role of volunteers in ecological observations and studies has increased in the past decade, as the scientific investment in ecology and long-running surveys by scientific institutions have decreased. Museums and other public engagement with science organisations are also active in citizen science, and it is part of their mission to educate and engage people with science. Researchers in universities and similar institutions engage with citizen science partly because it offers access to new resources that they would not be able to access otherwise without public support – for example computing resources in the Search for Extra Terrestrial Intelligence in the SETI@Home project. In some cases, such as large-scale classification tasks, there might be some similarities with the practice of crowdsourcing. Another issue within citizen science is access to the datasets that result from the efforts of volunteers'. Here, there is a growing practice to share the results back with volunteers or the wider scientific community, to ensure that volunteers are credited in publications as seen in efforts such as the Global Biodiversity Information Facility (GBIF[8]) which was set up to streamline the sharing of open and free biodiversity data.

Community science, as a form if citizen science, is of special interest, as it provides an example where new forms of factual spatial data are created and used to progress community goals. The Public Laboratory of Open Technology and Science,[9] illustrates this. Public Lab is a community of environmental activists and technology experts that promotes the use of low-cost adapted ('hacked') technology to monitor environmental issues (Dosemagen, Warren and Wylie, 2011). One of their early efforts was the creation of an aerial imagery apparatus using a kite or balloon to carry a cheap digital camera to take a large set of images over a relatively small area. The images that the camera captured are then sorted and stitched together to create a continuous image over the area where the balloon or kite was flown. This large-scale imagery provides visible evidence that is then annotated with additional information to highlight community issues – for example to provide evidence on how many participate in a demonstration, or the impact of a new road on a Palestinian village in Jerusalem.[10] In Public Lab work, affordable technology is combined with community's effort to provide instructions and guidance, which in turn supports the efforts to inform a situation of local concern. In such situations, citizen science is a tool of empowerment in the political sense, as it provides 'hard evidence' that emerges from scientific instruments or sensing devices, and methodology which supports a specific narrative that is of importance to the people who put it forward. It becomes an accepted form of evidence based decision making.

13.4 Development and Future Directions

Even though VGI and citizen science have much longer histories, most of the attention from policy makers, researchers and businesses has only been expressed in the past

[8] http://www.gbif.org/
[9] http://publiclab.org/
[10] http://publiclab.org/wiki/jerusalem

decade. Frequently, questions arise about the quality of the resulting information as well as the motivation of the participants (see Sieber & Haklay 2015). As more evidence emerges to confirm the quality of the data and that participants' motivations is recognized as not being mass recruitment exercise, attention can turn toward the compilation of longitudinal data collection. In some VGI activities the collected information is 'hyper-local' – making it only relevant to a small area in both space and time – for example information about a traffic jam and its implication on navigational decisions. Yet, even this localised information has relevance at a wider scale. In most VGI datasets, and especially in the area of citizen science, there is a need to understand how the information changes over time. Thus, the activities in these fields have the duality of describing a snapshot of the world (capturing an observation at a specific time and place and recording it), yet, because of the continuous sharing of the information, the dataset as a whole is always dynamic and in a state of change (see Perkins 2014 for detailed analysis of the (im)mutability of OpenStreetMap).

The process in which the information is produced, controlled and shared demonstrates differences in the power relationship and in financial benefits as discussed earlier. Concurrently, the ability to maintain the repository of information over time should receive more attention. For example, OpenStreetMap servers require regular operating system updates and effort as well as resources to deal with hardware failures. Sustainability requires an organisation, institution or a company to take responsibility. As a result, the control of the system (understood here in the wider sense and not just the hardware/software part of it) foregrounds issues of power, control and resources into these seemingly distributed, non-hierarchical activities. In addition, the process of data quality assurance requires oversight and moderation of more experienced and knowledgeable participants who check the information provided by novices. Over time, power relationships reveal themselves in both the case of VGI and citizen science.

13.5 Conclusion

As VGI and citizen science activities progress, the questions regarding data quality, the longevity of engagement, incentives and motivation of volunteers as well as the nature of the types of participants persist. Some research examining differential power differences has begun (Sieber & Haklay 2015), there is nontheless plenty of scope to critically study VGI and citizen science. For example, there are different levels of inclusiveness in terms of who is involved in data collection and which areas are being monitored. There is also merit to further investigate organisational practices and cultural influence with regards to the recruitment and ability to retain the participants over time. Aspects of gender inequality has are being discussed (Cooper & Smith, 2010, Stephens 2013), while ethnic, socioeconomic and age have received less attention. There is also scope to understand how wider politics and economic incentives lead to outcomes; for example, which thematic areas receive attention and funding and, more specifically, how did the production of base maps be perceived as a valuable commercial activity, while the recording of biodiversity is not? By understanding VGI and citizen science as a socio-technical system, and giving due attention to social aspects, might provide better insight about the nature of the spatial data being produced through these activities and what these tell us about the state of the world.

References

Becker, M., Caminiti, S., Fiorella, D., Francis, L., Gravino, P., Haklay, M., Hotho, A., Loreto, V., Mueller, J., Ricchiuti, F., Servedio, V. DP., Sîrbu, A., & Tria F. (2013). Awareness and Learning in Participatory Noise Sensing. *PLoS One* **8**(12):e81638.

Brown, J. (2001). Three Case Studies, in Werry C., & Mowbray, M. (eds) *Online communities: commerce, community action, and the virtual university*, Upper Saddle River, NJ: Prentice Hall.

Calabrese Barton, A. M. (2012). Citizen(s') Science: A Response to "The Future of Citizen Science". *Democracy and Education*, **20**(2), Article 12. Available at: http://democracyeducationjournal.org/home/vol20/iss2/12

Cooper, C. B. (2012). Links and Distinctions Among Citizenship, Science, and Citizen Science: A Response to "The Future of Citizen Science". *Democracy and Education*, **20**(2), Article 13.
Available at: http://democracyeducationjournal.org/home/vol20/iss2/13

Cooper, C. B., & Smith, J. A. (2010). Gender patterns in bird-related recreation in the USA and UK. *Ecology and Society*, 15(4), 4.

Craglia, M., Ostermann, F., & Spinsanti, L. (2012). Digital Earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, **5**(5), 398-416.

Dosemagen, S., Warren, J., & Wylie, S. (2011). Grassroots mapping: Creating a participatory map-making process centered on discourse. *Journal of Aesthetics and Protest*, *8*.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, **69**(4), 211–221.

Haklay, M. (2013). Citizen Science and Volunteered Geographic Information – overview and typology of participation in Sui, D.Z., Elwood, S. & M.F. Goodchild (eds.), 2013. Crowdsourcing Geographic Knowledge. Berlin: Springer. pp. 105-122.

Haklay, M. (2015). *Citizen Science and Policy: a European Perspective*. Woodrow Wilson Center for International Scholars, Washington DC.

Haklay, M., Singleton, A., & Parker, C. (2008). Web mapping 2.0: the Neogeography of the Geoweb. *Geography Compass, 3* 2011-2039

Howe, J. (2006). The Rise of Crowdsourcing. *Wired Magazine*. June 2006.

Laurel, B. (2001). *Utopian entrepreneur*. MIT press.

Mueller, M. P., Tippins, D., & Bryan, L. A. (2012). The Future of Citizen Science. *Democracy and Education*, **20**(1), Article 2.
Available at: http://democracyeducationjournal.org/home/vol20/iss1/2

Oxford English Dictionary (OED) (2014). *Citizen Science*, viewed August 2014.

Perkins, C. (2014). Plotting practices and politics: (Im)mutable narratives in OpenStreetMap. *Transactions of the Institute of British Geographers*, **39**(2), 304-317.

Shirky, C., (2011). *Cognitive Surplus: Creativity and Generosity in a Connected Age*, Penguin.

Sieber, R. E., & Haklay, M. (2015). The epistemology (s) of volunteered geographic information: a critique. *Geo: Geography and Environment* 2.2 (2015): 122-136.

Silverman, J. (2014). The Crowdsourcing Scam: Why do you deceive yourself? *The Baffler* (26). Available at http://www.thebaffler.com/salvos/crowdsourcing-scam accessed April 2015

Stephens, M. (2013). Gender and the GeoWeb: divisions in the production of user-generated cartographic information. *GeoJournal*, 78(6), 981-996.