Geographical Citizen Science – clash of cultures and new opportunities

Introduction

When considering the role of Volunteer Geographical Information ('VGI' coined by Goodchild 2007) in advancing science, it is necessary to place this type of VGI in its correct context. For example VGI can be seen as a way of producing geographical information, and as a tool for updating national geographical databases (Antoniou, Haklay & Morley in press) in which case the appropriate context is spatial data quality and the production of geographical information. In the current paper, the appropriate context for the evaluation of VGI's role in advancing science is the area of Citizen Science. Citizen science can be defined as scientific activities in which non-professional scientists participate in data collection, analysis and dissemination of a scientific project (Cohn 2008; Silvertown 2009). Notice that, by definition, citizen science can only exist in a world in which science is socially constructed as the preserve of professional scientists in academic institutions and industry, because, otherwise, any person who is involved in a scientific project would simply be considered as a contributor and potentially as a scientist. As Silvertown (2009) noted, until the late 19th century, science was mainly developed by people who had additional sources of employment that allowed them to spend time on data collection and analysis. Famously, Charles Darwin joined the *Beagle* voyage not as a professional naturalist, but as a companion to Captain FitzRoy. Thus, in that era, all science was citizen science.

Even with the rise of the professional scientist, the role of volunteers has not disappeared, especially in areas such as archaeology, where it is common for enthusiasts to join expeditions, or in natural science and ecology, where they collected and sent samples and observations to national repositories. These activities include the Christmas Bird Watch that has been ongoing since 1900 and the British Trust for Ornithology Survey, which has collected over 31 million records since its establishment in 1932 (Silvertown 2009).

While this type of citizen science has occurred in specific enclaves of scientific practice, the emergence of the Internet and the Web as a global infrastructure has enabled a new incarnation of this phenomena, with the proliferation of information and communication technology, the realisation of scientists that the public can provide free labour, skills, computing power and even funding, and finally the growing demands from research funders for public engagement all contributing to the motivation of scientists to develop and launch new citizen science projects (Silvertown 2009; Cohn 2008).

Citizen science can take a form in which volunteers put their efforts to a purely scientific endeavour, such as mapping galaxies, or a different form that might be termed 'community science' in which scientific measurements and analysis are carried out by members of local communities so they can develop an evidence base and set action plans to deal with problems in their area. A successful example of such an approach is the 'Global Community Monitor' method to allow communities to deal with air pollution issues (see gcmonitor.org). This is performed through a simple method of sampling air using plastic buckets followed by analysis in an air pollution laboratory, with, finally, the community provided with instructions on how to understand the results. This activity was termed

'Bucket Brigade' and was used across the world in environmental justice campaigns. In London, community science was used to collect noise readings in two communities that are impacted by airport and industrial activities. The outputs were effective in bringing environmental problems to the policy arena (Haklay, Francis & Whitaker 2008).

A specific sub-type of citizen science activities can be termed 'Geographical Citizen Science', and includes projects where the collection of location information is an integral part of the activity. A lot of the information that was and is collected through citizen science projects is geographical – as in the location of observations in the Christmas Bird Watch, or the sighting of a specific botanic species in ecological projects. Yet, in the past the location was usually approximate, and sometimes given in grid co-ordinates with 100-metre, or even 1 kilometre, accuracy, which meant that linking observations to a location could be tricky and highly uncertain. The changes that ushered in the Web Mapping 2.0 era (Haklay, Singleton & Parker 2008) are at the basis of accurate location information that can be obtained with little effort through the use of technology such as GPS receivers. This type of data collection belongs to the same class of activities that Goodchild (2007) identified. Even though location technology is now becoming increasingly available through Personal Navigation Devices (PND), GPS receivers and mobile phones, it is important to remember that, because of their affordance (in the sense of familiarity and ease of use, as in Norman 1990), paper maps are a very effective medium for data collection. This has been shown in citizen science studies in London as well as in the integration of paper mapping activities in OpenStreetMap through Walking Papers (Migurski 2009). As long as the data collection media supports accurate geographic location, the project can result in high-quality geographical citizen science.

The rest of the paper is dedicated to a short sketch of the likely characteristics of geographical citizen science and to one of the most significant challenges that it faces: overcoming the current culture of science.

Geographical Citizen Science – typical characteristics

The first important characteristic of geographical citizen science is to understand the role of the volunteer. The role can be active or passive. An active contribution happens when the participant is expected to consciously and actively contribute to the observation or the analysis, as in the case of taking an image of an observed species, tagging it and sending it electronically to the project's hub. A passive data collection can happen when the contributor is acting more as an observation platform and the data is gathered without active engagement. For example, when a person volunteers to be tagged by a GPS receiver to monitor daily walking activity or when a volunteer replaces a memory card and batteries in an automatic digital camera that is installed on a deer track (Cohn 2008). A further differentiation can be made between geographically explicit and implicit citizen science projects (Antoniou, Haklay & Morley in press). In geographically explicit projects, the activity is aimed at collecting geographical information; for example, in the British Trust of Ornithology project where a participant is required to record a specific location where the observation is taking place. In geographically implicit projects, the aim might be to collect images of different species, and some of the images will arrive geotagged, but the aim of the project is not to collect geographical information.

These different schemes will have an impact on the need to motivate and engage the participants and the level of training and knowledge that are required from her. In addition, they will influence the ability to secure quantitative and qualitative information. While all forms can support the collection of some quantitative information, only the active and geographically explicit applications are likely to provide meaningful qualitative information such as descriptions of personal perceptions or textual description of the place where the observation was recorded.

Research into the motivation (Budhahthoki, Nedovic-Budic & Bruce 2010; Coleman, Georgiadou & Labonte 2009) and the spatial characteristics of VGI (Haklay in press), as well as studies of Wikipedia and citizen science projects, is providing some guidance on what it is possible to achieve with geographical citizen science. What we know is that citizen science is a 'serious leisure' activity and that the most likely participants will join with some existing interest in the subject, and will be keen to learn more. They will be predominantly male, well educated and from higher brackets of the income scale, which gives them both the time capital to participate in the activity and the financial resources for specialist equipment and/or participation in field work. There will be 'participation inequality' with some participants contributing a lot, while many other contribute a little.

The areas that can be covered well by geographical citizen science will be areas with a high population concentration or high level of outdoor activity such as popular national parks. Coverage of other areas requires special planning and creation of suitable motivational schemes, including monetary ones or reliance on a smaller pool of volunteers. There will also be temporal aspects to data collection, with summer months, weekends and daytime being the more popular times.

All this means that the data is essentially heterogeneous, and it is important to remember that the data quality will vary according to the number of volunteers who work on the data and the particular knowledge of each volunteer. Thus, the data should not be treated as homogeneous and complete where a quality statement can be attached to the whole assemblage but, rather, localised measures must be used.

While all these aspects of geographical citizen science have an impact on the type of research questions and scientific challenges that can be answered through utilisation of suitable schemes, there is a potentially more significant challenge hindering this use of citizen science. This issue is more to do with the contemporary culture of science than with the putative value of citizen science itself.

Culture of Science problems

The story of modern science is often told by highlighting the increased precision and accuracy with which information is obtained and analysed (for example, Bryson 2004). In experimental science the instruments and devices were designed over the years to provide better accuracy, and complex experimental protocols were created to ensure that the level of uncertainty associated with measurements was reduced.

While it is well recognised that different academic disciplines have their own culture and specific practices (Latour 1993), such as the practice of double blind studies in medicine but not in other areas that research human subjects, the issue of dealing with uncertainty has been central to many areas of science, including geographical information science (see Couclelis 2003).

Interestingly, the attempt to eliminate uncertainty is especially prevalent in areas where science is used in practical applications such as engineering or environmental management. Due to organisational reasons and policy (King & Kraemer 1993), protocols are enshrined in regulations and become 'the correct way of doing science' with rigid protocols, some parts of which are arbitrary.

The strive to eradicate uncertainty (or at least reduce it) and the development of complex protocols are at the heart of the cultural issue that is leading to suspicion, derision and dismissal of geographical citizen science as a valuable method of scientific data collection.

The mistrust of geographical citizen science is, as noted in the opening, based on the view that science is best left to scientists, and it requires rigour, knowledge and skills that only professional scientists develop over time. As Silvertown (2009) noted: 'The apparent underrepresentation of citizen science in the formal literature probably has two causes. First, the term itself is relatively recent, and in fact hundreds of scientific papers have resulted from the data collected in Christmas Bird Counts and other long-running volunteer monitoring programmes. Second, **projects that fit uneasily into the standard model of hypothesis-testing research are written about only in the grey literature, or even not at all**;' (p. 471, emphasis added). This is also highlighted in Holling (1998) who emphasised that there are two cultures of science, and that citizen science by necessity belongs to the type of science that incorporates uncertainty and highlights integrative approaches.

Interestingly, suspicion of VGI follows along the lines of what Holling identified in the area of ecology over a decade ago. Despite the evidence that VGI can be as accurate as professional data (Haklay in press; Girres and Touya in press), mistrust of VGI as a useful source of geographical information is common among professional users (for example, Flanagin & Metzger 2008).

What is forgotten by those who oppose geographical citizen science is the development of instrumentation and its impact on the balance of knowledge and skills that are required by the operator, as well as the level of motivation, dedication and attention to detail of volunteers.

Scientific instrumentation has evolved tremendously over the past 350 years, if we take the invention of instruments such as Hooke's microscope in 1663 as an indication of the development of modern scientific instruments. Since then, the instruments have improved tremendously in terms of their observational power, their accuracy and precision. However, until fairly late in the 20th century, they demanded an increasing amount of knowledge in order to operate them effectively. A great deal of professional judgement was required to balance the accuracy of a calculation with the practical aspects of conducting research. Consider, for example, that many calculations for the moon missions were carried out with slide rules, where experience and judgement is necessary to decide if the calculations are satisfactory.

The computerisation and miniaturisation of scientific instruments, especially in the past 20 years, have changed this equation. Now, the humble GPS receiver encapsulates knowledge and procedures that are highly complex and is capable of calculating them with good precision without any intervention from the user. The GPS satellites themselves also encapsulate significant amounts of scientific knowledge and understanding. For example, William Roy spent about 6 weeks and engaged a team of 12 men to measure the 5-mile baseline of the English triangulation system with an accuracy of 2.5 inches. Today, a single person, equipped with a standard GPS receiver and a mobile phone can achieve a similar fit in less than a day. The sophistication of the equipment and, more importantly, the science that is integrated into the computational parts of it, enables this.

The ability of equipment to provide accurate and precise measurement is central to the ability of volunteers to provide reliable scientific information, especially when these are used in tandem with their personal knowledge and commitment. For example, research in the US has shown that citizen scientists identified crab types correctly 95% of the time (Cohn 2008). Citizen scientists show significant commitment to the topic and are as capable as the best researchers, in many cases. Thus, the information that they produce should be trusted.

Summary

Geographical citizen science has clearly grown in recent years and is showing significant potential in areas such as biodiversity, air pollution or recording the changing shapes of cities. There are, however, two issues that are critical when considering its role in advancing science.

First and foremost, there is a need to consider which scientific questions can be answered by it according to the patterns of data collection, the ability to recruit and train volunteers, and other aspects of VGI. Second, there is a need to overcome the cultural issue and to develop an understanding and acceptance of citizen science within the scientific community. This will require challenging some of the deeply held views in science, such as viewing uncertainty not as something that can be eliminated through tighter protocols but as an integral part of any data collection, and therefore developing appropriate methods to deal with it during analysis.

One intriguing possibility is that citizen science will work as an integral part of participatory science in which the whole scientific process is performed in collaboration with the wider public. Some examples are already emerging in geography (Pain 2004) and might provide directions for the future development of citizen science projects.

References

Antoniou, B., Haklay, M., and Morley, J., in press. Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon, *Geomatica, The Journal of Geospatial Information, Technology and Practice* (special issue on VGI)

Bryson, B., 2004. *A short history of almost everything*, Black Swan

Budhathoki, N. R., Nedovic-Budic, Z., and Bruce, B., 2010. An Interdisciplinary Frame for Understanding Volunteered Geographic Information. *Geomatica, The Journal of Geospatial Information, Technology and Practice*, 64(1)

Cohn, J. P., 2008. Citizen science: Can volunteers do real research? *Bioscience* 58(3):192–197

Coleman, D.J., Georgiadou, Y., and Labonte, J., 2009. Volunteered geographic information: the nature and motivation of producers. *International Journal of Spatial Data Infrastructures Research* 4: 332-358

Couclelis, H., 2003. The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS* 7(2):165-75

Flanagin, A.J., and Metzger, M.J., 2008. The credibility of volunteered geographic information. *GeoJournal* 72(3-4):137-148

Girres, J-F., and Touya, G., in press. Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*

Goodchild, M., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal* 69:211–21

Haklay, M., Francis, L., and Whitaker, C., 2008. Citizens tackle noise pollution, *GIS professional*, August issue

Haklay, M., in press. How good is OpenStreetMap information? A comparative study of OpenStreetMap and Ordnance Survey datasets for London and the rest of England *Environment and Planning B*

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

Holling, C.S., 1998. Two cultures of ecology. *Conservation Ecology* [online] 2(2): 4. Available WWW http://www.consecol.org/vol2/iss2/art4/

King, J. L., and Kraemer L. K., 1993. Models, facts, and the policy process: the political ecology of estimated truth. In *Environmental modeling with GIS*. Eds. Michael F. Goodchild, Bradley O. Parks, and Louis T. Steyaert, 353-60. New York: Oxford University Press

Latour, B., 1993. *We have never been modern*, Harvard University Press

Migurski, M., 2009. *Walking Papers* (walking-papers.org)

Norman, D. A., 1990. *The Design of Everyday Things*. New York: Doubleday

Pain, R., 2004. Social Geography: Participatory Research. *Progress in Human Geography* 28(5):652-663

Silvertown, J., 2009. A new dawn for Citizen Science. *Trends in Ecology & Evolution* 24(9):467-471