Exploring the Global Adoption of Citizen Science

Ramine Tinati
University of Southampton
Web and Internet Science
r.tinati@ecs.soton.ac.uk

Markus Luczak-Roesch
University of Southampton
Web and Internet Science
mlr1m12@soton.ac.uk

Elena Simperl
University of Southampton
Web and Internet Science
elena.simperl@soton.ac.uk

Wendy Hall
University of Southampton
Web and Internet Science
wh@soton.ac.uk

ABSTRACT

In recent years there has been a growing interest toward the application of Web-based citizen science platforms. Such platforms use crowdsourcing techniques to support scientific advancements, and in several cases, have lead to new scientific discoveries which were not originally considered. Our work explores the highly successful Web-based citizen science platform, Zooniverse, a crowdsourcing platform with a userbase of over 1 million participants who volunteer their free time to support scientific enquiries. We focus on the growth of the Zooniverse platform, which has evolved from a rudimentary crowdsourcing platform where users were presented with tasks, into a platform which has become a rich community of citizen scientists, discussion, and interaction. Building upon existing research into the motivations and design considerations of developing and sustaining citizen science projects, this paper explores the space of citizen science engagement within the Zooniverse, and ask the question of whether citizen science has become a worldwide activity.

1. INTRODUCTION

In recent years, advances in science has greatly benefited from the growing interest and engagement with crowdsourcing platforms. This growth in so called *citizen science* platforms have managed to successfully attract and engaged millions of unpaid volunteers, to spend their time, answering science-related tasks, such as the classification of objects within millions of images, or the transcription of ancient manuscripts and logs. Expanding from the early citizen science astronomy-related projects, scientists from multiple domains and disciplines are now using crowdsourcing techniques in order to support and advance their areas of research. Consequently, projects, such as the Zooniverse, are developing platforms to support an eco-system of projects, which provides volunteers with a broad range of projects to engage with.

Moreover, citizen science as an activity has evolved, platforms such as the Zooniverse have grown from a solo activity where the entire user workflow would consist of answering the pre-defined set of questions or tasks, to an environment where volunteers discuss, and collaborate about their observations, and – often serendipitously – discover new scientific knowledge.

However, whilst citizen science has shown to help various domains, project success is not guaranteed, which has been the driver to several studies of citizen science participation, from exploring their motivations, to modelling their behaviour to reduce churn whilst increasing accuracy [14]. This has also led to the analysis and review of citizen science design and development, with research documenting the design processes, and principles of building these special kinds of crowdsourcing platforms [16].

Building on the literature concerning citizen science, our work explores the space of understanding the adoption of citizen science as a global activity. We wish to understand how citizen science is perceived beyond the cultural borders of the western world, and see if the uptake of such an activity is a global phenomenon, or something that is biased towards societies that are more familiar to an ethos of volunteering. In this paper we discuss our previous work conducted on citizen science motivations and design, and then present preliminary findings related to the adoption of citizen science as a worldwide activity. We then consider these findings in terms of the wider cultural and societal nuances that may influence adoption.

2. RELATED WORK

Online citizen science systems have typically enlisted the help of participants with three kinds of activities: data collection/curation (e.g., [19]); data analysis (e.g. [17, 6]); and problem-solving (e.g., [8, 2]). The Zooniverse [5] fits the second category; small professional science teams, who provide the raw digital artefacts for each project, such as images, video and audio recordings, are supported by vast numbers of members participants who voluntarily identify, classify, and label the data according to instructions and loose direction provided by the science teams.

Previous work has studied a variety of different aspects of citizen science systems. Analysis of the factors that most influence accuracy and completion of tasks, as well as methods for improving such performance such as through refined task selection have been studied [1, 7], as well as task scheduling heuristics [8]. Wiggens et al. [18] analysed CS projects as online communities and modelled CS community organisation through the relationship between volunteering, digital technologies, and workplaces. Other studies have sought to understand the core motivations of citizen scientists (e.g. [11, 12]), while others have focused on documenting the

increasingly common citizen-initiated serendipitous scientific discoveries [13]. Such studies often reference the vast psychology and social psychology literature relating to human motivation to provide a theoretical grounding for continued engagement with CS systems [4].

2.1 Summary of Previous Work

In summary, our previous work has explored several areas of interest within the Zooniverse and other citizen science platforms. We have explored motivations of citizen scientists and their behaviour characteristics [14], as well as a detail analysis of the discussion and collaborative activity between citizen scientists, and how this results in the discovery of unknown scientific finding [10]. The use of gamification as a mechanism for citizen science design was also investigated, specifically looking at how certain features can be used to improve user retention and engagement [15]. Drawing upon these findings, and more widely within the field of online community research, a design framework for citizen science projects was developed, which describes the various social and technical features which may improve the chances of launching a successful citizen science project [16].

3. DATA AND RESULTS
3.1 Dataset

Our study draws upon 5 months of data from the Zooniverse platform collected between May to September 2015. which contains $7,746,840$ classification and $120,009$ talk records, from $61,833$ users. The classification and talk records contain various fields of information about the logged in user, including their identifier, the timestamp, project and item they were working on, and the message text if they were posting on the messaging forum. Classifications also contain the geographic location of the logged in user, which is given by a latitude and longitude coordinates of where the classification was performed.

3.2 Results

We conducted a series of descriptive statistical tests on the dataset which provided us with a macro-level perspective of the volunteers engaging within the task feature, and also within the talk integrated talk forums within Zooniverse. As Table 3.2 shows, a significant proportion of the total number of classifications were conducted by members identified as being located in Europe or North America, cumulatively accounting for 88% of the total number of classifications.

Table 3.2 provides a list of the top 10 languages identified within either the task or talk records (corresponding the preferred UI language, or the body of the message published in the forum). As shown, English accounted for 86.3% of the total entries, which corresponds to the previous analysis of where Zooniverse users are geographically located.

4. DISCUSSION

In existing studies of citizen science, or more broadly in online communities research, the analysis of user participation and engagement have often asked questions about *why* people take part. Studies are often framed by motivational frameworks, surveys and qualitative studies, seek to answer questions about the intrinsic and extrinsic desires of people willing to contribute (either for free or for paid). Findings often reveal a series of motivating factors which can more than often be synthesised down to a set of basic human needs; altruism, power, money, love, or a sense of self-belonging [9].

Region	Classifications
Europe (UK, Germany, France)	3688453 (48.2%)
North America (USA, Canada, Mexico)	3071134 (40.2%)
Oceania (Australia, New Zealand, Tanzania)	347818 (4.6%)
Asia (Singapore, India, Japan)	277536 (3.6%)
Far East	*37278 (0.5%)*
Middle East	*15318 (0.2%)*
South America (Brazil, Argentina, Chile)	154807 (2.0%)
Africa (South Africa, Egypt, Kenya)	50045 (0.7%)

Table 1: Classifications made aggregated by geographical region (Asia count is a cumulative value which includes the Far East and Middle East). For each region and the corresponding top 3 contributing countries are listed

Language	Users (%)
English	86.3
French	3.3
German	2.1
Spanish	1.4
Polish	1.0
Chuvash	1.0
Czech	1.0
Russian	0.9
Dutch	0.8
Italian	0.7

Table 2: Top 10 identified languages of users for talk or task activity. Language for the task entry is based on the preferred language of the UI, chosen by the user. In total 38 languages were identified

However, as the initial analysis has shown, there is a significant proportion of the world that are *not* taking part in these activities.

The preliminary findings raise several questions to both the designers of citizen science platform, and those interested in understanding citizen science as a new form of Web activity. For a citizen science platform design and community manager, these findings have multiple implications from the projects design, interface, features, and languages available, to the domain (e.g. nature, or Astronomy) of the project. Furthermore, if citizen science aims to become a globally accessible and successful activity, then understanding the reasons for why specific regions are not engaged, is required.

The results suggest that the Zooniverse platform has managed to engaged significantly within the Western world, and is predominately driven by an English speaking community. However, whilst containing over half of the worlds population, countries within regions such as Asian and Africa contribute to a very small proportion of the contributions made. Intuitively, there are several societal and economic factors that limit the participation of people from countries within these regions, however there are several countries within these regions which are in the socio-economic position to contribute.

In a bid to answer such questions, and unpack the external factors affect the level participation observed, our future work will turn towards the social sciences to explore the role of cultural norms and practises, with specific interest in the concept of volunteering and its acceptance within society [3]. We will draw upon theories of work and volunteering culture in order to understand the deeper reasons for why participant is lower in such areas, and what mechanisms, social and technical, can be employed by citizen science developers to improve engagement.

5. REFERENCES

[1] Cooper, C. Zen in the art of citizen science: Apps for collective discovery and the 1 percent rule of the web. `http://blogs.scientificamerican.com/guest-blog/2013/09/11/zen-in-the-art-of-citizen-science-apps-for-collective-discovery-and-the-1-rule-of-the-web/`, 2013.

[2] Cordero, P., Lucks, J. B., and Das, R. An rna mapping database for curating rna structure mapping experiments. *Bioinformatics 28*, 22 (2012), 3006–3008.

[3] Dekker, P., and Halman, L. *The values of volunteering: Cross-cultural perspectives*. Springer Science & Business Media, 2003.

[4] Ebner, W., Leimeister, J. M., and Krcmar, H. Community engineering for innovations: the ideas competition as a method to nurture a virtual community for innovations. *R&D Management 39*, 4 (2009), 342–356.

[5] Fortson, L., Masters, K., Nichol, R., Borne, K., Edmondson, E., Lintott, C., Raddick, J., Schawinski, K., and Wallin, J. Galaxy zoo: Morphological classification and citizen science. *arXiv preprint arXiv:1104.5513* (2011).

[6] Heinzelman, J., and Waters, C. Crowdsourcing crisis information in disaster-affected haiti. `http://www.humanitarianforum.org/data/files/crowdsourcinghaiti.pdf`, 2010.

[7] Kawrykow, A., Roumanis, G., Kam, A., Kwak, D., Leung, C., Wu, C., Zarour, E., Sarmenta, L., Blanchette, M., Waldispühl, J., et al. Phylo: a citizen science approach for improving multiple sequence alignment. *PloS one 7*, 3 (2012), e31362.

[8] Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popović, Z., Baker, D., and Players, F. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences 108*, 47 (2011), 18949–18953.

[9] Lampe, C., Wash, R., Velasquez, A., and Ozkaya, E. Motivations to participate in online communities. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2010), 1927–1936.

[10] Luczak-Rösch, M., Tinati, R., Simperl, E., Kleek, M. V., Shadbolt, N., and Simpson, R. Why won't aliens talk to us? content and community dynamics in online citizen science. In *Eighth International AAAI Conference on Weblogs and Social Media* (2014).

[11] Raddick, J., Lintott, C., Bamford, S., Land, K., Locksmith, D., Murray, P., Nichol, B., Schawinski, K., Slosar, A., Szalay, A., et al. Galaxy zoo: Motivations of citizen scientists. In *Bulletin of the American Astronomical Society*, vol. 40 (2008), 240.

[12] Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D., and Jacobs, D. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM (2012), 217–226.

[13] Schwamb, M. E., Orosz, J. A., Carter, J. A., Welsh, W. F., Fischer, D. A., Torres, G., Howard, A. W., Crepp, J. R., Keel, W. C., Lintott, C. J., et al. Planet hunters: A transiting circumbinary planet in a quadruple star system. *The Astrophysical Journal 768*, 2 (2013), 127.

[14] Tinati, R., Luczak-Roesch, M., Simperl, E., and Shadbolt, N. Motivations of citizen scientists: A quantitative investigation of forum participation. In *Proceedings of the 2014 ACM Conference on Web Science*, WebSci '14, ACM (New York, NY, USA, 2014), 295–296.

[15] Tinati, R., Luczak-Rösch, M., Simperl, E., Hall, W., and Shadbolt, N. /command'and conquer: analysing discussion in a citizen science game. *ACM Web Science Conference 2015* (2015).

[16] Tinati, R., Van Kleek, M., Simperl, E., Luczak-Rösch, M., Simpson, R., and Shadbolt, N. Designing for citizen data analysis: A cross-sectional case study of a multi-domain citizen science platform. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, ACM (New York, NY, USA, 2015), 4069–4078.

[17] Westphal, A. J., Butterworth, A. L., Snead, C. J., Craig, N., Anderson, D., Jones, S. M., Brownlee, D. E., Farnsworth, R., and Zolensky, M. E. Stardust@ home: a massively distributed public search for interstellar dust in the stardust interstellar dust collector.

[18] Wiggins, A., and Crowston, K. From conservation to crowdsourcing: A typology of citizen science. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, IEEE (2011), 1–10.

[19] Zook, M., Graham, M., Shelton, T., and Gorman, S. Volunteered geographic information and crowdsourcing disaster relief: a case study of the haitian earthquake. *World Medical & Health Policy 2*, 2 (2010), 7–33.