Creating user profiles based on citizen scientists' engagement patterns

Van Galen, Coen
coen.vangalen@ru.nl
Radboud University, Netherlands, The

van Oort, Thunnis
thunnisvanoort@yahoo.co.uk
Radboud University, Netherlands, The

Prats López, Montserrat
montserrat.pratslopez@ou.nl
Open Universiteit, The Netherlands

Wessel, Ganzevoort
wessel.ganzevoort@ru.nl
Radboud University, Netherlands, The

Mourits, Rick
rick.mourits@iisg.nl
International Institute of Social History, the Netherlands

Crowdsourcing is increasingly being used in scientific research projects of different disciplines and hence often referred to as crowd science or online citizen science. This new way of conducting research entails the online participation of citizens in research projects usually initiated by professional scientists. Capturing the interest of citizens to participate in a citizen science project and keeping them engaged for the duration of a (sometimes lengthy) project are important and challenging activities for project organizers (West & Pateman, 2016).

The citizen science literature provides recommendations for recruitment and engagement strategies and the most important one is to take into account the different motivations of citizen participants and use engagement strategies accordingly (West & Pateman, 2016; Ponciano & Brasilerio, 2014). However, we argue that knowing why people decide to participate is not enough to design effective engagement strategies. Recent research proposes to focus on citizen scientists' online behaviours or actions (Ponciano & Brasileiro, 2014; Jackson et al., 2016; Aristeidou et al., 2017). An examination of the patterns of actual participants' behaviour can be used to distinguish between various ways of engaging in a project (De Moor et al., 2019). Knowledge about engagement patterns based on citizen scientists' behaviour is a valuable input for project coordinators to improve citizens' experience and continued involvement in a project (Aristeidou et al., 2017).

To study these engagement patterns we work as a multidisciplinary team, bringing together expertise from various domains including history and information sciences. The aim of our research is to identify different types of engagement in an online humanities project, assess the differences and similarities with research carried out in other disciplines and provide recommendations for project coordinators. To this purpose, we work within the citizen science project 'Historical Database Suriname and Curaçao' that includes over 650 participants (https://hdsc.ning.com/; van Galen, 2019). The objective of this citizen science project is to create a database of the population of Suriname and Curaçao from the years 1830 to 1950, by digitizing and transcribing civil registers and death certificates. This database will be open access available for both scientists and the public in general, to facilitate genealogical research, the study of social processes and diversity in colonial society as well as the repercussions of slavery over multiple generations.

For our analyses of participant engagement, we use the log data of the first 5 months of this citizen science project. These data indicate when (pseudonymized) users performed transcription tasks or posted messages on the forum. Based on the log data, we calculated for each user the number of entries, the number of sessions (defined as a series of uninterrupted entries or with a gap smaller than 60 minutes), the number of entry days and the active period per user. These data were in turn used to measure users' activity ratio and variation in periodicity (Aristeidou et al., 2017). The log data do not provide information on the duration of a task, therefore, task duration is approximated with the timestamps between finished tasks. To measure the daily devoted time (Ponciano & Brasileiro, 2014), the session duration per day was used and missing information filled in with the average task duration across all users. This set of indices per user were the basis to distinguish clusters (k-means) of users.

The log data and the measures described above allows us to discern patterns of how users engage with the tasks, forming a basis to create user profiles. For instance, we have identified a group of users who perform tasks regularly throughout a period ("regulars") versus users that show sudden bursts of activity ("sprinters"); or users that are active on the forum but less active in transcribing (and the other way around) – suggesting user types with different kinds of gratification derived from their activity as citizen scientists .

To help us interpret the meaning of the various user profiles, we will have in-depth interviews with a sample of users (and former users). Both the results of the analysis of the empirical log data combined with the coded interviews will provide a better understanding of how to differentiate between various types of online citizen scientists, each with their own patterns of engagement, allowing project leaders to tailor efforts to recruit, motivate and engage participants, and provide a better experience.

Bibliography

Aristeidou, M. / **Scanlon, E.** / **Sharples, M.** (2017) : "Profiles of engagement in online communities of citizen science participation", Computers in Human Behavior, 74, 246-256

De Moor, T / **et al.** (2019): "Dynamics of Engagement in Citizen Science: Results from the "Yes, I do!"-Project", *Citizen Science: Theory and Practice*, 4, 1: 1–17. DOI: https://doi.org/10.5334/cstp.212

Jackson, C. B. / **Østerlund, C.** / **Maidel, V.** / **Crowston, K.** / **Mugar, G.** (2016): "Which Way Did They Go?: Newcomer Movement through the Zooniverse", *Proceedings of the 19th CSCW Conference*, San Francisco, California, USA.

Ponciano, L. / **Brasileiro, F.** (2014): "Finding Volunteers' Engagement Profiles in Human Computation for Citizen Science Projects", *Human Computation*, 1, 2: 247-266.

Van Galen, C. W. (2019): "Creating an audience: Experiences from the Surinamese slave registers crowdsourcing project", *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 52, 3: 178-194, DOI: 10.1080/01615440.2019.1590268

West, S. / **Pateman, R.** (2016): “Recruiting and Retaining Participants in Citizen Science: What Can Be Learned from the Volunteering Literature?“, *Citizen Science: Theory and Practice*, *1*, 2: 15.