JCOM

Invisible brokers: "citizen science" on Twitter

Elise Tancoigne

Abstract
Who speaks for "citizen science" on Twitter? Which territory of citizen science have they made visible so far? This paper offers the first description of the community of users who dedicate their online social media identity to citizen science. It shows that Twitter users who identify with the term "citizen science" are mostly U.S. science professionals in environmental sciences, and rarely projects' participants. In contrast to the original concept of "citizen science", defined as a direct relationship between scientists and lay participants, this paper makes visible a third category of individual actors, mostly women, who connect these lay participants and scientists: the "citizen science broker".

Context
What does a Scandinavian cook watching birds have in common with a Spanish student measuring quality of air in her neighbourhood, an American engineer giving spare computer power to calculating projects, and a Mexican patient sharing her health data online?[1] Today these individuals are clustered under the term "citizen science participants" or "citizen scientists." The term "citizen science" is commonly attributed to the ornithologist Rick Bonney at the Cornell laboratory of Ornithology, U.S.A., who has popularized it since the mid-1990s to refer to public participation in the collection of observations for scientists [Bonney, 1996]. In 2014, the Oxford English Dictionary defined the word in the following way: *"Citizen science n. Scientific work undertaken by members of the general public, often in collaboration with or under the direction of professional scientists and scientific institutions."* [Oxford English Dictionary Online, 2014].

The public has participated in producing scientific knowledge for a long time, in ways that have not been restricted to this type of "citizen science" [Strasser et al.,

[1]For more information on these projects, see: eBird (https://ebird.org/), Air Quality Citizen Science (https://aqcitizenscience.rti.org/), SETI@home (https://setiathome.berkeley.edu/), and PatientsLikeMe (https://www.patientslikeme.com/).

2019]. The term captures a wider diversity of approaches and encompasses historical cases such as nineteenth-century amateur naturalists studying birds migration in Germany [Mahr, 2014] and the participation of French patient organizations in biomedical research in the 1980s [Epstein, 1995]. Long before the term "citizen science" was coined, the Brazilian educator Paulo Freire inspired many to engage in "community based (action) research" or "participatory action research" [Freire, 2000], connecting researchers and lay people in projects to produce knowledge that could help solve local problems [Gutberlet, Tremblay and Moraes, 2014]. Today some of these older types of participatory research have been eclipsed by projects carried out under the more popular heading of "citizen science." SciStarter, a U.S. web platform created in 2009, assigns the term to over 1,600 projects [SciStarter, 2019], including invitations to the public to help cure HIV/AIDS [Foldit, 2018], develop infrared cameras [Public Lab, 2018] and make communities resilient to climate change [CSC-ATL, 2018].

As the number of "citizen science" projects have grown and generated more interest, this transformation in the research landscape has caught the attention of scholars who have proposed a number of typologies and definitions to better grasp the extent and significance of the phenomenon [Eitzel et al., 2017; Lewenstein, 2016]. In 2015, Follett and Strezov [2015] analysed 1,127 Scopus and Web of Science articles indexed as "citizen science" and documented a prevalence of biological and biodiversity studies, particularly in ornithology. Two further scientometric studies based on Web of Science [Kullenberg and Kasperowski, 2016; Cointet and Joly, 2016] confirmed the importance of environmental research carried out under the label, while noting a significant public participation in geographic information research, epidemiology [Kullenberg and Kasperowski, 2016], and agricultural research [Cointet and Joly, 2016] categorized under other headings.

Other studies have gone beyond the question of the nature of the scientific disciplines involved. Pettibone, Vohland and Ziegler [2017] examined the institutions behind 97 projects from two major German-language "citizen science" platforms (the German *buergerschaffenwissen.de* and the Austrian *citizen-science.at*). They showed that the key actors were research organizations, followed by so-called "society-based groups" (non-profit organizations, independent groups), government agencies and media organisations. The study confirmed the prevalence of participatory research devoted to biodiversity and environmental monitoring. The same year, Pocock et al. [2017] analysed 509 web environmental and ecological participatory projects through a range of variables (protocols, supporting resources, data accessibility, modes of communication and project scale), but provided little additional information on the individual and organizations involved in these projects.

Admittedly, many projects made surveys of their 'lay participants'.
Some data collected on the demographic characteristics of the participants show that they are mainly white, male, and middle-aged [Curtis, 2015; Reed et al., 2013; Raddick, Bracey, Gay, Lintott, Murray et al., 2010]. They are mostly described in terms of motivations, thought processes [Trumbull et al., 2000; Price and Lee, 2013], and science literacy [Jordan et al., 2011; Crall et al., 2013; Brossard, Lewenstein and Bonney, 2005], qualities that matter most in attracting or retaining participants, and measuring project outcomes. Lack of information is even more acute with regard to the project organizers. A few studies have explored scientists' attitudes towards par-

ticipatory research [Riesch and Potter, 2014; Golumbic et al., 2017], but an extensive socio-demographic study of those who organize citizen science projects is lacking.

Furthermore, all these approaches have tended to reinforce the idea that "citizen science" consists solely of a relationship between professional scientists and lay people (or 'citizens'). They overlook who might be the "brokers" of citizen science, i.e. "people or organizations that move knowledge around and create connections between researchers and their various audiences" [Meyer, 2010]. Two features of "citizen science" that are frequently emphasized, indeed suggest that a number of brokers play a crucial role in citizen science. When seen as a way to educate people about science [Bonney, 1996], one should expect an involvement of science teachers, science journalists, outreach professionals and educational institutions. Secondly, to the extent that citizen science heavily relies on digital platforms such as Scistarter (a third of whose projects exist exclusively online) or Zooniverse, one can expect computer scientists, developers, designers, and engineers to play important roles as well.

Objective

Rather than offering another depiction of the field through scientometrics or survey approaches, this article offers an original attempt at identifying the spokespersons of citizen science, through their presence on a widely used social media: Twitter. Who are the advocates of citizen science? Are they researchers, participants, brokers? To which territory of citizen science do they belong? Which territory have they made visible so far?

The public social media Twitter offers a particularly promising proxy to identify the various advocates of participatory research under the banner "citizen science". Indeed, Facebook and Twitter have been shown to offer a particularly rich source for studying participant engagement [Robson et al., 2013] and citizen science project managers make an extensive use of diverse social media [Liberatore et al., 2018], particularly for recruiting large numbers of participants. Moreover, social media include both individuals and organizations, a key asset to understand their respective roles in advocating for citizen sciences.

Methods

Working with Twitter

Launched in 2006, Twitter is a public social media used by almost one quarter of the U.S. adult population and predominantly by young adults: in the U.S., half of Twitter users are adults in their twenties, a third are in their thirties or forties, and a fifth are in their fifties or sixties [Pew Research Center, 2019]. Data available for retrieval include users' tweets, biographies and followers-followees links. Tweets are any text, video or image messages users post on Twitter. They may also contain hashtags (keywords starting with "#"), URL links, or mentions to other users (names starting with "@"). Twitter has given users the possibility to create profiles in a 160-character section for biographical sketches or 'bios'. Most Twitter users define themselves on the basis of their occupation, interests, or role they play in society ("I am a Gamer, a YouTuber"), and often their relationships ("Happily married to @John") [Priante et al., 2016]. Followers and followees' relationships are connections made by a user to another user. They are asymmetrical: one may

follow a user but not be followed back. Retweets, mentions in tweets and followers-followees' relationship are often the basis of Twitter sub-communities studies [Abdelsadek et al., 2018].

Identifying actors of citizen science on Twitter

Community detection in Twitter often starts with retrieving accounts sharing specific keywords in their tweets or bio. This paper follows the choice made by Grandjean [2016] and Wang et al. [2017] to work on users' biographies, instead of tweets, for three main reasons. First, Twitter's free API limits keyword search in tweets to only a sample of the tweets published in the last 7 days. On the other hand, access to users' biographies is not limited in time and includes both active and inactive users. Moreover, tweets may comment on or refer to other users' view, while *"users' personal profiles always pertain to themselves"* [Wang et al., 2017]. Users typically select quite carefully the traits which will be displayed to the audience [Bullingham and Vasconcelos, 2013], especially when the amount of information they can provide is tightly restricted, and the outcome is the product of negotiation between several facets of a user's identity. Actors with a strong commitment to "citizen science" are therefore expected to mention the term in their Twitter bios, while they may use synonyms of citizen science during 7 days of tweets. Another interesting point is the blurring of professions and hobbies in the Twitter bios, particularly when "creative professions" such as research are cited [Sloan et al., 2015]. This is a key point given that when the public participates in science, the boundary between work and leisure is often blurred [Stebbins, 1996]. The data retrieval process was kept voluntarily simple, so as to be easily reproducible — for instance in order to study the evolution of the citizen science's spokespersons on Twitter over time. The first step therefore consisted in retrieving the account data of the users whose bios contained the term "citizen science", "citizenscience", "citsci", or "citizen scientist" (with capital, plural, or singular) (October 2017) using the statistical software R and the package rtweet provided by Michael Kearney [2018].

Figure 1 summarises the process by which data from Twitter bios was obtained and analysed. The R script is provided in the supplementary material.

Qualifying actors of citizen science on Twitter

The second step consisted in dividing accounts into those that refer to an individual ("personal accounts") and those that refer to a collective (e.g. project, organization: "collective accounts") and doing attribute coding for all of them. Attribute coding is *"the notation [. . .] of basic descriptive information such as: the fieldwork setting (e.g., school name, city, country), participant characteristics or demographics (e.g., age, gender, ethnicity, health), data format (e.g., interview, field note, document), time frame and other variables of interest for qualitative and some applications of quantitative analysis. [. . .] It provides participant information and contexts for analysis and interpretation. [. . .] Virtually all qualitative studies will employ some form of Attribute Coding"* [Saldaña, 2012, pp. 55–56]. Instead of using surveys, whose response rate usually stands below 40% [Nov, Arazy and Anderson, 2014; Sheehan, 2001], users' basic demographic information was collected on the web. Most people use their real name, therefore information for 90% of the accounts could be derived from external sources such as LinkedIn, online CVs, or professional web pages. For
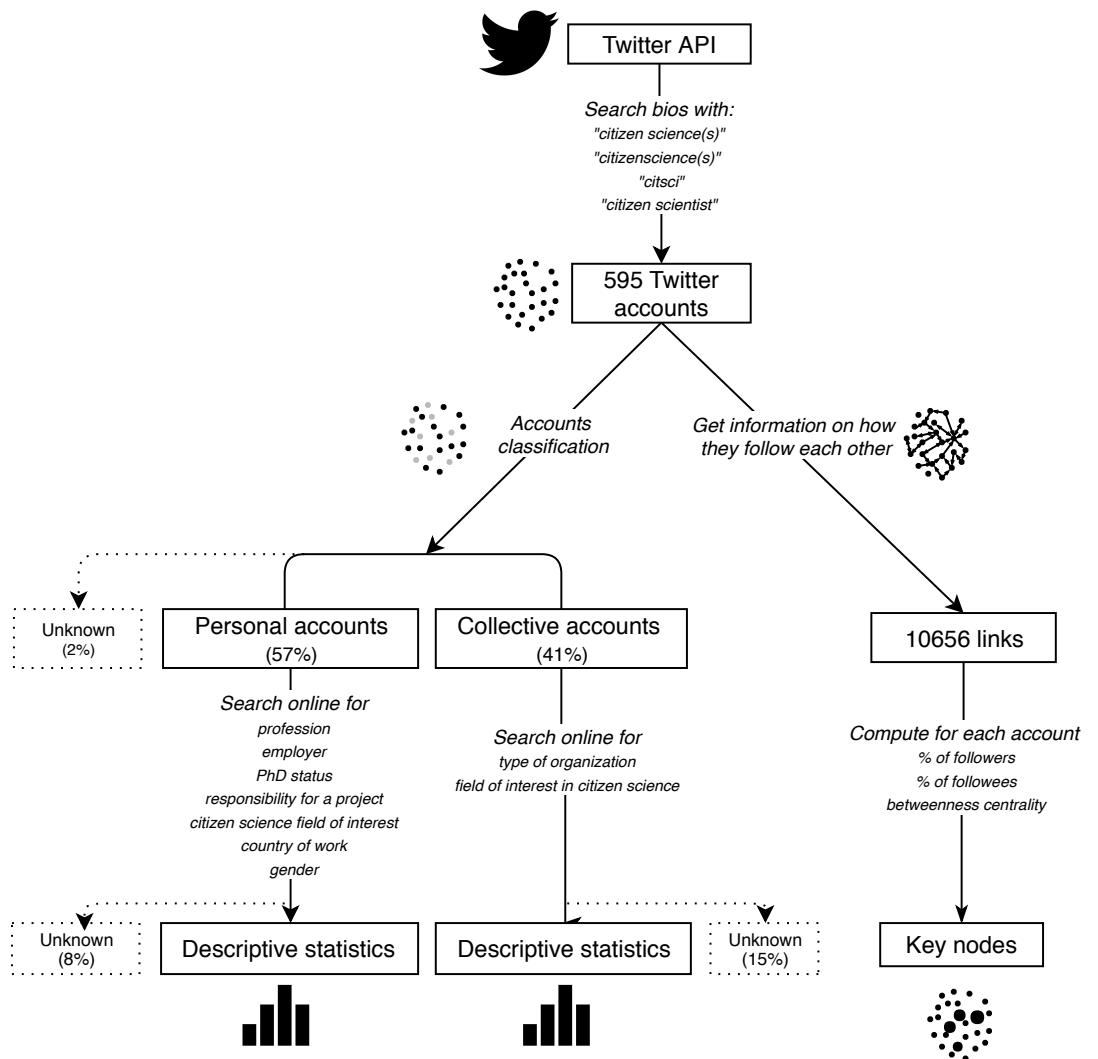
Figure 1. Process followed to gather and analyze data on actors that identify themselves with "citizen science" on Twitter.

personal accounts, information was collected on people's 1/*profession*, 2/*employer*, 3/*PhD status* (started/completed), 4/*responsibility for a project* (yes/no) and 5/*citizen science field of interest*. Variables 6/*gender* (male/female) and 7/*country of work* were also included so as to compare the "citizen science" population with the global population on Twitter. To characterize the collective accounts, online information was searched for the 1/*type of collective* and 2/the collective's *citizen science field of interest*. Table 1 provides descriptions and examples of this attribute coding process used for both personal and collective accounts. Codes were created in Open Office Calc, following a bottom-up process maximizing non-overlapping classification of raw data (see also the appendix, section "group codes"). The category "others" was chosen when no thematically coherent grouping could be found for the few accounts that remained. The results are displayed in Figure 2.

Table 1: Codes used to categorize the people and collectives that identify themselves with citizen science on Twitter.

Attribute	Code	Description of the code	Examples of raw data from LinkedIn and other sources
Personal accounts			
Profession	researcher	Graduate student or Ph.D. holder working in a research organization	*Adjunct lecturer, associate professor*
	technical staff	has a technical job	*engineer, data analyst*
	outreach professional	works as science journalist, writer, communicator, teacher, manager or coordinator for participatory projects	*nature guide, editor at [a science journal], Community Manager*
	other profession	employed in non-science related jobs	*jewellery creator, press officer*
Employer	research organization	universities and research institutes	*Roskilde University, The Scripps Research Institute*
	research-related organization	company, NGO or institute that promote, use or offer services to research organizations	*Rathenau Instituut, Silent Spring Institute, DigitalTech Consulting, Inc.*
	outreach organization	organization that present research to the public	*Newsday Media group, YCAM Bio Research, Swedish Museum of Natural History, Scofield Magnet Middle School, Baer Fishing Adventures, DNA Root-Search*
	government	organization related to the State	*Southampton City Council, U.S. Environmental Protection Agency*
	other	any other employer	*Kaiser Permanente, freelance psychotherapeut, Accenture*
	none	not engaged in a regularly paid activity	*stay-at-home mom*
Collective accounts			
Type of collective	citizen science projects	project that calls for participants. The category does not include projects whose aim is to better understand/promote participatory researches	*Open Air Lab, Blooms for Bees*
	structures which promote citizen science	organization specifically dedicated to the study or promotion of citizen science. May be a NGO, an online website	*SciStarter, DITOs*
	environmental NGO	includes NGOs dedicated to environment conservation and education	*Hawaii Nature Hui, ReefQuest*
	structures which popularize science	structure dedicated to the popularization of science. Includes NGOs/organisations; lessons, courses	*BetterBio, Science Cheerleaders*
	research teams or scientific societies	Research teams or scientific societies	*Conservation Gov Lab, Paleo Quest*
	companies	Companies selling products or services	*senseBox, The Martian Garden*

Continued on the next page.

Table 1: *Continued from the previous page.*

Attribute	Code	Description of the code	Examples of raw data from LinkedIn and other sources
Code used for both types of accounts			
Field of interest in citizen science	environment	related to conservation, biodiversity, ornithology, climate, environment at large	*HerpMapper, AuroraWatch-Net*
	astronomy	related to astronomy and space	*Ultrascope, NASA Solve*
	health	related to medicine, human body, diseases	*UT Biome Project, CiTIQUE*
	DIY	related to the making and building of objects	*Hackuarium, Biopalette, ReaGent*
	meta-citizen science	related to the study of participatory researches, from different perspectives: user experience, design, gaming, education, social sciences	*UX researcher, Learning Design specialist*
	other	related to other areas: linguistics, geography, urbanism, gender studies	*IT COUNTS, Lingscape*

Identifying key actors of citizen science on Twitter

Not everyone is equal on Twitter: some accounts are more active than others or have more influence than others. Providing a flat list of accounts therefore misses credit and visibility information within the field. On the other hand, getting the top accounts allows for in-depth analysis of the trajectories of the persons and collectives behind them (beyond gender or profession). Neither the number of followers nor the number of tweets are good estimates of an account's influence. A rich literature, mostly from the marketing and information sciences, has criticized these indicators and offered alternatives [Cha et al., 2010]. A combination of three network indicators based on the followers and followees relationships (see Figure 1) was used to identify key actors in the dataset. First, the proportion of "citizen science" accounts that each account attracts and connects to was computed with the igraph package in R [Csárdi and Nepusz, 2006]. For instance, an account followed by 30% of the other "citizen science" accounts, which follows only 2% of these accounts, has a more prominent place in the network than an account that follows a lot of accounts but is almost never followed. These measures are similar to the classical *authority* and *hub* scores which identify prominent sources of information within a network [Kleinberg, 1999]. The betweenness centrality [Freeman, 1977; Csárdi and Nepusz, 2006] of each node was then computed in order to identify accounts with a bridging position within the network. A node with a high betweenness centrality indicates that many paths in the network go through this node. Figure 3 presents the result in a two-dimensional plot, with key nodes displayed on the top-right corner. Finally, a fine-grained qualitative analysis was performed on these key nodes, searching for the employment history, memberships, and funding sources of the individuals or collectives they represent.

Ethical concerns   Appropriate institutional ethical review processes were followed. From a legal viewpoint, each user on Twitter signed the company's Terms of Service at

registration. These terms of service allow the use of Twitter data for research purpose without additional express opt-in consent from the users. This includes the practice of associating data from other sources with Twitter accounts (called "Off-Twitter matching"), provided these data come from public sources. Our analysis fits into these legal rules [Twitter, Inc., 2019]. However ethics is not just about complying with the law. Recently published guidelines emphasize the idea that the public dimension of data does not give researchers an unlimited right to use them [Townsend and Wallace, 2016] and also that *"each research situation is unique and it will not be possible simply to apply a standard template in order to guarantee ethical practice"* [British Sociological Association, 2017]. The present study ranks users on the basis of indicators that are not publicly available. It therefore produces information that is new and may cause discomfort to individuals or organizations concerned if they were identified. Williams, Burnap and Sloan [2017] indeed show that most Twitter users would like to be anonymized, should their tweets be cited. As bringing full people's and organization names does not bring anything to the analysis, they were fully anonymized in this publication.

Results

A predominance of science professionals

On Twitter, 336 personal accounts contain "citizen science" in their description. Researchers (PhD students, post-docs, or faculty) form the largest category in the dataset (36%, Figure 2A). Two-thirds of these researchers work with data produced by citizen science projects in various fields: environment (42%), health (9%), astronomy (8%) and other fields (10%). The last third consists of researchers who study participation in citizen science projects. These individuals work in the fields of user experience, design studies, gaming studies, science and technology studies, or science education.
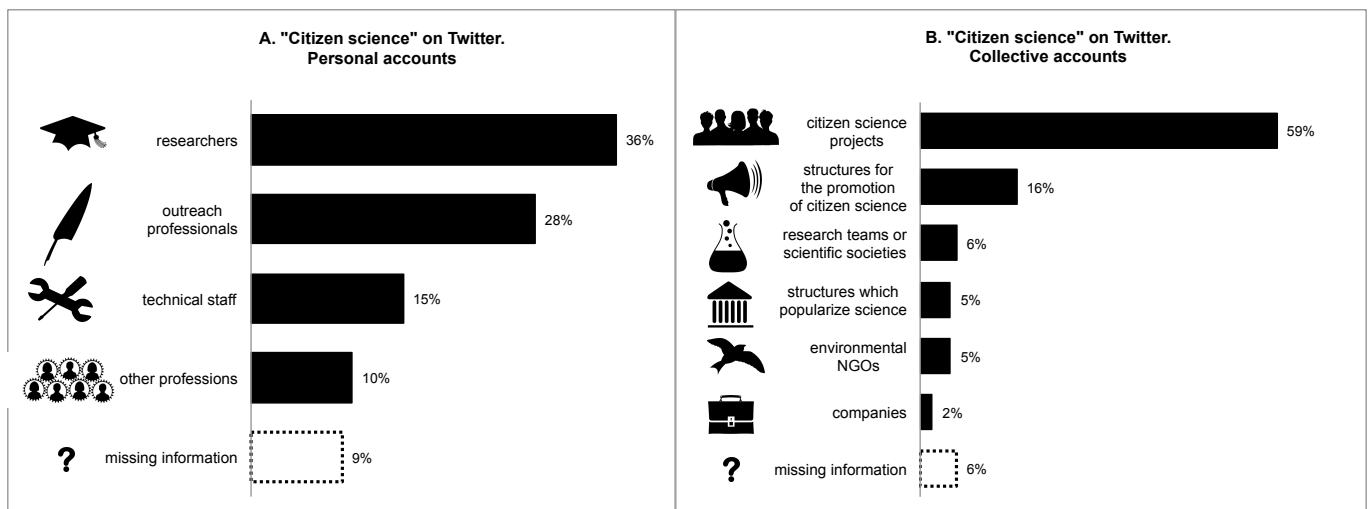

Figure 2. Description of the actors identifying with "citizen science" on Twitter. N = 336 (A), 268 (B).

The second most represented category of individual accounts includes people whose profession is related to public outreach (28%, Figure 2A): journalists, writers, teachers (15%) and project coordinators (13%). In addition to these professional

project coordinators, a minority of researchers take on this task as well, leaving that role predominantly to dedicated professionals.

Overall, over 80% of the individuals work in scientific research and outreach and 60% hold a PhD. Most work in public universities, research institutes, or museums (40%), NGOs, institutes and companies with research-related activities (20%), and a small number works in media and schools (7%).

The label "citizen *scientist*" is widely used in the media and project platforms to describe the lay participants. But only 16% of the people in this analysis designated themselves as "citizen scientists". Of those who did, two-thirds work in non-scientific fields (*jewellery creator* for instance) or remained silent on their occupation. The last third concerns people in outreach, and technical staff.

Dedicated organizations

Most of the collective accounts belong to specific citizen science projects (around 60%, see Figure 2B). Two-thirds of these are engaged in environmental projects (65%), a result consistent with the findings of Kullenberg and Kasperowski [2016] and Cointet and Joly [2016]. Second on the list are organizations more generally dedicated to the promotion of citizen science (Figure 3 shows some of the most prominent organizations in this category). At the bottom of the list are a few NGOs devoted to environmental science, science popularization organizations, research teams, scientific societies, or private companies. These actors do not consider "citizen science" a core aspect of their identities, thus are almost absent from the dataset.

Key actors identifying with citizen science: a small subset of U.S. actors funded by the NSF

Figure 3 shows the distribution of the "citizen science" accounts on Twitter according to their place in the network. Big dots close to the top right corner represent the most prominent nodes of the network: the key actors identifying with citizen science on Twitter.

These key actors may be collectives or individuals. The collectives include national associations dedicated to "citizen science", citizen science web resources and citizen science web media. The individuals are all women (see the discussion).

These individuals and collectives form a small subset of nodes related to each other through collaboration or membership. Many have a connection to the Cornell Laboratory of ornithology: one is a former PhD student, another a research associate and a third a public outreach professional of the Cornell Lab. The founder of one of the web resources and a public outreach professional working at the Natural History Museum of Los Angeles County are lacking a connection to the lab. All of these individuals play key roles in the two visible national citizen science associations. Three of these women were also advisors for Media #2 with one of them also managing Media #1.

Nearly all of these individuals and collectives have received funding from the National Science Foundation (NSF) [Strasser et al., 2019]. Such funding was awarded from the Division of Research on Learning for educational purposes (7M$
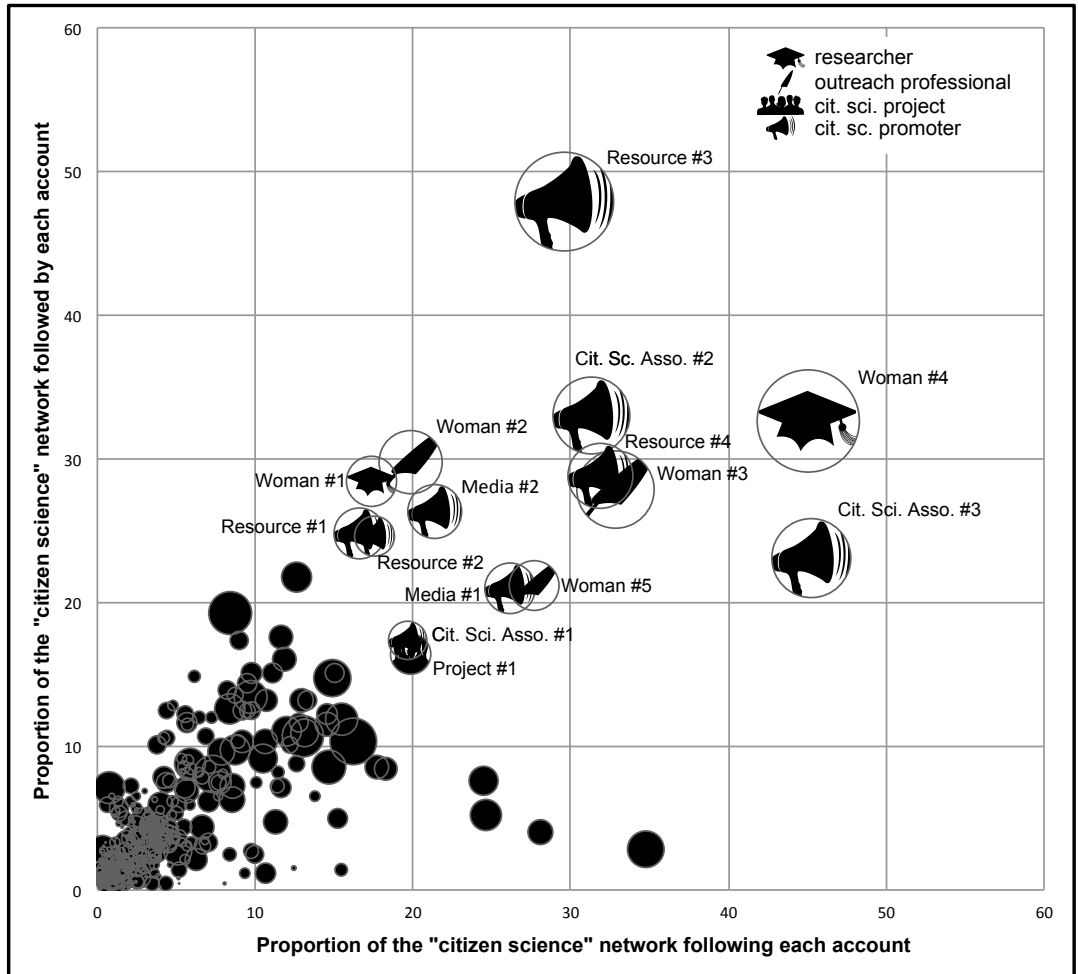
Figure 3. Key accounts identifying with "citizen science" on Twitter (N total=595). Nodes sized accorded to their betweenness centrality. Scale: Woman #4=28'291, Project #1=4'398. Cit. Sci. Asso.: Citizen Science Association.

in total), and for the promotion of research infrastructures from the Office of Advanced Cyberinfrastructure (2.5M$ in total). The awards permitted the Cornell Lab of Ornithology to create three of the top-organizations visible Figure 3, and Resource #4 at Colorado State University.

The predominance of American actors remains true beyond the key identified actors. In the dataset of all citizen science accounts, 46% are located in the United States, whereas only 22% of all Twitter accounts are located in the U.S. [Statista, 2018], confirming the current predominance of the United States in "citizen science".

Discussion

So then, which are the most visible proponents of citizen science today on the most important public social media? Accounts of individuals and collectives which refer to "citizen science" in their Twitter biographies reflect the makeup of the professional science community: researchers, technical staff, science outreach professionals and organizations. The key actors among them are based in the U.S., and are supported by governmental funding mainly for the benefit of

environmental projects for educational purposes. The predominance of people working at research organizations is consistent with the academic roots of the term [Bonney, 1996]. It is no surprise that citizen science projects are predominant on Twitter, since they use this social media to recruit participants and as a feedback infrastructure. Providing estimates of followers and tweets is a relatively simple way for project managers to generate the type of impact assessment that funders generally require. The predominance of environmental and U.S.-based actors is also a reflection of the ornithological, U.S. roots of the phenomenon.

Few "lay people" refer to "citizen science" in their Twitter biographies

Why does a very small number of "lay people" (people who are not science professionals) appear in the dataset, in contrast to the "millions of such participants" envisioned in the rhetoric surrounding citizen science [Bonney et al., 2014]? There are two possible explanations: the participants may not use Twitter, or do not identify themselves with "citizen science" when composing their Twitter bios.

Twitter is used predominantly by young adults, which contrasts with the demographics of volunteering in citizen science projects or other activities, online or otherwise. Statistics for the U.S. show that volunteers are predominantly in their thirties and fifties [Bureau of Labor Statistics, 2015] with those in their fifties doing so at even higher rates, often in areas such as conservation, environmental and animal rights organizations [Dolnicar and Randle, 2007]. The trend holds for those who volunteer online, as seen in the citizen science project Galaxy Zoo, with high participation among people in their fifties [n=10,708 respondents Raddick, Bracey, Gay, Lintott, Cardamone et al., 2013]. This indicates that profiles on Twitter are not the best method to identify volunteers, especially for environmental and online activities. Accounts of citizen science projects often report hundreds of followers (median= 484). uBiome, for instance, has more than 30,000 followers while Wildlife Watch has more than 10,000. Obviously, these are not all science professionals, which indicates that the second explanation might be true as well: many participants in citizen science projects with Twitter accounts do not use the labels "citizen science" or "citizen scientist" in their profiles. These are terms constantly used by organizers and the media, but even the top participants do not include the expression in their bios. For example, none of the top participants in three very different types of citizen science projects, iNaturalist (spotting fauna observations), Eyewire (data analysis to map the brain), and PatientsLikeMe (self-reporting) refer to themselves using the expression "citizen scientist" or "citizen science". Terry Hunefeld (@thunefeld), who ranked #2 on iNaturalist, states that he "Volunteers for the Anza Borrego Desert State Park. Loves birding." Neves Seraf (@NSeraf), ranked #1 on Eyewire, states: "I love cokiiies [sic] and Harry potter. I also like kardashians [sic]". These volunteers make no reference to science in their bios, but refer to a specific topical interest, project, or passion (like birding) instead. The participants themselves rarely take up this label, which was coined by science professionals. The creation of this label by scientists, with almost no citizen science coordinator or researcher using the term in her bio, resonates highly with the boundary making process described by Maney and Abraham [2008]: *"Boundary making involves creating symbolic distinctions between social groups. Who "we" are is defined in contradistinction to negatively represented others [Taylor and Whittier, 1992]. In a content analysis of focus group discussions about immigrants in Rotterdam,*

Verkuyten, de Jong, and Masson [1995] identify four forms of boundary making: labelling, metaphors, concretization, and commonplace". Labelling is part of creating "the other": by creating the category of "citizen scientists", the scientists go on drawing a line between them and the participants.

A high number of middle-range professionals: the brokers

In addition to learning that the expression "citizen science" is mainly used on Twitter by professionals who organize projects, we found that it is used by a high number of individuals who occupy middle-range between what the media call "real scientists" and "citizen scientists" or "lay participants" [SciStarter, 2019; Torpey, 2013; Lakshmanan, 2015]. This contrasts with definitions of citizen science that mention only "members of the general public" and "professional scientists" [Oxford English Dictionary Online, 2014]. Nearly 40% of the individuals in the dataset are outreach professionals, project coordinators and researchers who conduct research on citizen science. All these people are the brokers [Meyer, 2010] of citizen science: people who dedicate their efforts toward the development of sustainable connections between scientists and remote participants.

Why have the brokers of citizen science been a blind spot in the literature so far? One explanation might relate to this role being pretty new and accompanying the recent institutionalization of the field: all the national organisations were created after 2010. Or this may be understood as a rhetorical technology serving the recruitment of participants. It may be easier to sell "citizen science" to the public if there is an implicit promise that participants will "get in touch with 'real' scientists," since getting the opportunity to learn is one of the most commonly featured motivations of participants [Jackson et al., 2015; Jennett et al., 2016]. *"Meet the researchers who've created projects for free on the Zooniverse"* [Zooniverse, 2018] is probably more appealing than "Meet the coordinators and designers behind the infrastructure for this project." Promising a straight relationship between researchers and participants may therefore appeal to potential learners. A third explanation relates to a striking feature of these brokers, namely that about 70% are women. This figure contrasts starkly with overall estimates of gender distribution in science outreach, research or other science professions. In 2013, women constituted only 29% of those with an occupation in the fields of science & engineering [National Science Foundation, 2016], and 29% of the total of researchers worldwide [UNESCO Institute for Statistics (UIS), 2018]. The proportion of women journalists in the dataset (83%) also largely exceeds the proportion observed in the U.S.A. (55%) or Europe (45%) [SciDev, 2013, p. 2]. The skew in our figures cannot be explained by a bias in the Twitter audience, either; the platform has a similar penetration rate among men and women [Pew Research Center, 2019]. However, the most recent report on gender biases in S&T from the European Commission argues that [European Commission, 2016] *"In the higher education sector in most countries, men are more likely than women to be employed as researchers, whereas women are more likely than men to be employed as other supporting staff or technicians."* Moreover, women are more likely to be involved into community management functions (61% of community managers were women in 2013) [Keath, 2013] or tasks related to social media when involved in a collective. The high proportion of women in the dataset therefore reflects the more general tendency of women to occupy outreach functions for science organizations. Since

science has a long history of making women's work invisible [Abbate, 2017], it should come as no surprise that women occupy brokers' positions, known to be invisibilized [Meyer, 2010].

Importance of dedicated organizations

Finally, these results stress the importance of organizations and projects in the promotion of "citizen science". Recently organizations have been created that are specifically dedicated to the promotion of public participation in research, generally through governmental funding slated for educational or infrastructural purposes; they play a role in the advancement and circulation of knowledge on citizen science [Roche and Davis, 2017; Storksdieck et al., 2016].

Conclusion

Overall, this study shows that the community identifying with "citizen science" present on Twitter has very specific socio-demographic characteristics. The term "citizen science" is mostly used by professionals working at research and outreach institutions, particularly in environmental research. Most of the key actors of this network constitute a small and well-connected world with links to the Cornell Lab of Ornithology, where the term was coined. The citizen science model visible on Twitter is quite specific as it mostly centres around calls for volunteers to take part in predefined academic research projects.

Another interesting point relates to the high number of brokers involved in recruiting and creating knowledge around the participation of 'lay people' in the production of scientific knowledge. While the role of community organizers and support staff is also emphasized and recognized as a crucial facilitating factor in community-based research [Israel et al., 1998], this is not yet the case within the citizen science community. Characterizing the role played by these brokers within these projects in terms of how they frame science, society, as well as their own role, would be a further step in our understanding of the making of the citizen science industry.

Given the high proportion of women involved in such roles, it would be particularly interesting to know if their specific experience, as women, and their "situated knowledge" [Haraway, 2003] influence the design of projects and research questions.

Much has been written on whether participatory research should qualify as proper science or not [Cohn, 2008]. As the corpus of literature on the engineering of participation in non-scientific fields grows, bridges can be built to better understand the situation of citizen science [Gregory and Atkins, 2018] and the roles play by its brokers in it. It would be interesting to shift the focus from the comparison between citizen science and professional science to a comparison between scientific and non-scientific forms of volunteering, in terms of participation engineering and the role social media play in this engineering.

References　Abbate, J. (2017). Recoding gender: women's changing participation in computing. Cambridge, MA, U.S.A.: MIT Press Ltd.

Abdelsadek, Y., Chelghoum, K., Herrmann, F., Kacem, I. and Otjacques, B. (2018). 'Community extraction and visualization in social networks applied to Twitter'. *Information Sciences* 424, pp. 204–223. https://doi.org/10.1016/j.ins.2017.09.022.

Bonney, R. (1996). 'Citizen science. A lab tradition'. *Living Bird: For the Study and Conservation of Birds* 15 (4), pp. 7–15.

Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J. and Parrish, J. K. (2014). 'Citizen science. Next steps for citizen science'. *Science* 343 (6178), pp. 1436–1437. https://doi.org/10.1126/science.1251554. PMID: 24675940.

British Sociological Association (2017). *Ethics guidelines and collated resources for digital research*. URL: https://www.britsoc.co.uk/media/24309/bsa_statement_of_ethical_practice_annexe.pdf.

Brossard, D., Lewenstein, B. and Bonney, R. (2005). 'Scientific knowledge and attitude change: The impact of a citizen science project'. *International Journal of Science Education* 27 (9), pp. 1099–1121. https://doi.org/10.1080/09500690500069483.

Bullingham, L. and Vasconcelos, A. C. (2013). ''The presentation of self in the online world': Goffman and the study of online identities'. *Journal of Information Science* 39 (1), pp. 101–112. https://doi.org/10.1177/0165551512470051.

Bureau of Labor Statistics (2015). *Volunteering in the United States*. URL: https://www.bls.gov/news.release/volun.toc.htm.

Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, P. K. (2010). 'Measuring user influence in Twitter: the million follower fallacy'. In: *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (ICWSM)*, pp. 10–17.

Cohn, J. P. (2008). 'Citizen Science: Can Volunteers Do Real Research?' *BioScience* 58 (3), pp. 192–197. https://doi.org/10.1641/B580303.

Cointet, J.-P. and Joly, P.-B. (2016). 'Annexe 3. Analyse scientométrique des publications sur les sciences participatives'. In: Les sciences participatives en France. État des lieux, bonnes pratiques & recommandations. Rapport pour les ministres de l'education et de la recherche. Paris, France.

Crall, A. W., Jordan, R., Holfelder, K., Newman, G. J., Graham, J. and Waller, D. M. (2013). 'The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy'. *Public Understanding of Science* 22 (6), pp. 745–764. https://doi.org/10.1177/0963662511434894. PMID: 23825234.

Csárdi, G. and Nepusz, T. (2006). 'The igraph software package for complex network research'. *InterJournal Complex Systems* 1695 (5), pp. 1–9.
URL: http://www.interjournal.org/manuscript_abstract.php?361100992.

CSC-ATL (2018). *Center for Sustainable Communities*. URL: https://csc-atl.org/ (visited on 19th July 2019).

Curtis, V. (2015). 'Motivation to Participate in an Online Citizen Science Game: a Study of Foldit'. *Science Communication* 37 (6), pp. 723–746.
https://doi.org/10.1177/1075547015609322.

Dolnicar, S. and Randle, M. (2007). 'The international volunteering market: market segments and competitive relations'. *International Journal of Nonprofit and Voluntary Sector Marketing* 12 (4), pp. 350–370.
https://doi.org/10.1002/nvsm.292.

Eitzel, M. V., Cappadonna, J. L., Santos-Lang, C., Duerr, R. E., Virapongse, A., West, S. E., Kyba, C. C. M., Bowser, A., Cooper, C. B., Sforzi, A., Metcalfe, A. N., Harris, E. S., Thiel, M., Haklay, M., Ponciano, L., Roche, J., Ceccaroni, L., Shilling, F. M., Dörler, D., Heigl, F., Kiessling, T., Davis, B. Y. and Jiang, Q. (2017). 'Citizen Science Terminology Matters: Exploring Key Terms'. *Citizen Science: Theory and Practice* 2 (1), pp. 1–20. https://doi.org/10.5334/cstp.96.

Epstein, S. (1995). 'The Construction of Lay Expertise: AIDS Activism and the Forging of Credibility in the Reform of Clinical Trials'. *Science, Technology & Human Values* 20 (4), pp. 408–437.
https://doi.org/10.1177/016224399502000402.

European Commission (2016). SHE Figures 2015. Brussels, Belgium: Directorate-General for Research and Innovation. URL: https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf.

Foldit (2018). *Solve puzzles for science — Foldit*.
URL: https://fold.it/portal/index.php?q= (visited on 19th July 2019).

Follett, R. and Strezov, V. (2015). 'An analysis of citizen science based research: usage and publication patterns'. *PLOS ONE* 10 (11), e0143687.
https://doi.org/10.1371/journal.pone.0143687.

Freeman, L. C. (1977). 'A set of measures of centrality based on betweenness'. *Sociometry* 40 (1), p. 35. https://doi.org/10.2307/3033543.

Freire, P. (2000). Pedagogy of the oppressed. 30th anniversary ed. New York, NY, U.S.A.: Continuum.

Golumbic, Y. N., Orr, D., Baram-Tsabari, A. and Fishbain, B. (2017). 'Between vision and reality: a case study of scientists' views on citizen science'. *Citizen Science: Theory and Practice* 2 (1), p. 6. https://doi.org/10.5334/cstp.53.

Grandjean, M. (2016). 'A social network analysis of Twitter: mapping the digital humanities community'. *Cogent Arts & Humanities* 3 (1).
https://doi.org/10.1080/23311983.2016.1171458.

Gregory, A. J. and Atkins, J. P. (2018). 'Community operational research and citizen science: two icons in need of each other?' *European Journal of Operational Research* 268 (3), pp. 1111–1124. https://doi.org/10.1016/j.ejor.2017.07.037.

Gutberlet, J., Tremblay, C. and Moraes, C. (2014). 'The community-based research tradition in Latin America'. In: Higher education and community-based research. Ed. by R. Munck, L. McIlrath, B. Hall and R. Tandon. New York, NY, U.S.A.: Palgrave Macmillan U.S., pp. 167–180.
https://doi.org/10.1057/9781137385284_12.

Haraway, D. (2003). 'Situated knowledges: the science question in feminism and the privilege of partial perspective'. In: Turning points in qualitative research: tying knots in a handkerchief. Ed. by Y. S. Lincoln and N. K. Denzin. Lanham, MD, U.S.A.: Rowman Altamira.

Israel, B. A., Schulz, A. J., Parker, E. A. and Becker, A. B. (1998). 'Review of community-based research: assessing partnership approaches to improve public health'. *Annual Review of Public Health* 19, pp. 173–202. https://doi.org/10.1146/annurev.publhealth.19.1.173.

Jackson, C. B., Østerlund, C., Mugar, G., Hassman, K. D. and Crowston, K. (2015). 'Motivations for sustained participation in crowdsourcing: case studies of citizen science on the role of talk'. In: *2015 48th Hawaii International Conference on System Sciences*. IEEE, pp. 1624–1634. https://doi.org/10.1109/hicss.2015.196.

Jennett, C., Kloetzer, L., Schneider, D., Iacovides, I., Cox, A., Gold, M., Fuchs, B., Eveleigh, A., Mathieu, K., Ajani, Z. and Talsi, Y. (2016). 'Motivations, learning and creativity in online citizen science'. *JCOM* 15 (03), A05. https://doi.org/10.22323/2.15030205.

Jordan, R. C., Gray, S. A., Howe, D. V., Brooks, W. R. and Ehrenfeld, J. G. (2011). 'Knowledge Gain and Behavioral Change in Citizen-Science Programs: Citizen-Scientist Knowledge Gain'. *Conservation Biology* 25 (6), pp. 1148–1154. https://doi.org/10.1111/j.1523-1739.2011.01745.x.

Kearney, M. W. (2018). *Rtweet: collecting Twitter data*. (version 0.6.7). [R package]. URL: https://cran.r-project.org/package=rtweet.

Keath, J. (2013). *The 2013 community manager report*. [Infographic]. URL: https://www.socialfresh.com/community-manager-report-2013/.

Kleinberg, J. M. (1999). 'Authoritative sources in a hyperlinked environment'. *Journal of the ACM* 46 (5), pp. 604–632. https://doi.org/10.1145/324133.324140.

Kullenberg, C. and Kasperowski, D. (2016). 'What is Citizen Science? A Scientometric Meta-Analysis'. *Plos One* 11 (1), e0147152. https://doi.org/10.1371/journal.pone.0147152.

Lakshmanan, I. (2015). *The environmental outlook: citizen scientists*. URL: https://dianerehm.org/shows/2015-05-05/the-environmental-outlook-citizen-scientists.

Lewenstein, B. V. (2016). 'Can we understand citizen science?' *JCOM* 15 (01), E, pp. 1–5. URL: https://jcom.sissa.it/archive/15/01/JCOM_1501_2016_E.

Liberatore, A., Bowkett, E., MacLeod, C. J., Spurr, E. and Longnecker, N. (2018). 'Social media as a platform for a citizen science community of practice'. *Citizen Science: Theory and Practice* 3 (1). https://doi.org/10.5334/cstp.108.

Mahr, D. (2014). Citizen Science: partizipative Wissenschaft im späten 19. und frühen 20. Jahrhundert. Baden-Baden, Deutschland: Nomos.

Maney, G. M. and Abraham, M. (2008). 'Whose backyard? Boundary making in NIMBY opposition to immigrant services'. *Social Justice, Migrant Labor and Contested Public Space* 35 (4), pp. 66–82. URL: https://www.jstor.org/stable/29768515.

Meyer, M. (2010). 'The Rise of the Knowledge Broker'. *Science Communication* 32 (1), pp. 118–127. https://doi.org/10.1177/1075547009359797.

National Science Foundation (2016). '3. Science and engineering labor force'. In: Science and engineering indicators 2016. U.S.A.: National Science Foundation, p. 124.

Nov, O., Arazy, O. and Anderson, D. (2014). 'ScientistsHome: what drives the quantity and quality of online citizen science participation?' *PloS One* 9 (4), e90375. https://doi.org/10.1371/journal.pone.0090375.

Oxford English Dictionary Online (2014). *Citizen Science, n.*
URL: https://www.oed.com/view/Entry/33513#eid316619124 (visited on 19th July 2019).

Pettibone, L., Vohland, K. and Ziegler, D. (2017). 'Understanding the (inter)disciplinary and institutional diversity of citizen science: a survey of current practice in Germany and Austria'. *PLoS ONE* 12 (6), e0178778. https://doi.org/10.1371/journal.pone.0178778.

Pew Research Center (12th June 2019). *Social media fact sheet*.
URL: http://www.pewinternet.org/fact-sheet/social-media/.

Pocock, M. J. O., Tweddle, J. C., Savage, J., Robinson, L. D. and Roy, H. E. (2017). 'The diversity and evolution of ecological and environmental citizen science'. *PLOS ONE* 12 (4), e0172579.
https://doi.org/10.1371/journal.pone.0172579.

Priante, A., Hiemstra, D., van den Broek, T., Saeed, A., Ehrenhard, M. and Need, A. (2016). '#WhoAmI in 160 characters? Classifying social identities based on Twitter profile descriptions'. In: *Proceedings of the First Workshop on NLP and Computational Social Science*. Association for Computational Linguistics.
https://doi.org/10.18653/v1/w16-5608.

Price, C. A. and Lee, H.-S. (2013). 'Changes in participants' scientific attitudes and epistemological beliefs during an astronomical citizen science project'. *Journal of Research in Science Teaching* 50 (7). https://doi.org/10.1002/tea.21090.

Public Lab (2018). *Public lab: near-infrared camera*.
URL: https://publiclab.org/wiki/near-infrared-camera (visited on 19th July 2019).

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Cardamone, C., Murray, P., Schawinski, K., Szalay, A. S. and Vandenberg, J. (2013). 'Galaxy Zoo: Motivations of Citizen Scientists'. *Astronomy Education Review* 12 (1), pp. 010106–010101. https://doi.org/10.3847/AER2011021. arXiv: 1303.6886.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., Szalay, A. S. and Vandenberg, J. (2010). 'Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers'. *Astronomy Education Review* 9 (1), 010103, pp. 1–18. https://doi.org/10.3847/AER2009036. arXiv: 0909.2925.

Reed, J., Raddick, M. J., Lardner, A. and Carney, K. (2013). 'An Exploratory Factor Analysis of Motivations for Participating in Zooniverse, a Collection of Virtual Citizen Science Projects'. In: *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013)* (7th–10th January 2013). IEEE, pp. 610–619. https://doi.org/10.1109/HICSS.2013.85.

Riesch, H. and Potter, C. (2014). 'Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions'. *Public Understanding of Science* 23 (1), pp. 107–120. https://doi.org/10.1177/0963662513497324.

Robson, C., Hearst, M., Kau, C. and Pierce, J. (2013). 'Comparing the use of social networking and traditional media channels for promoting citizen science'. In: *Proceedings of the 2013 conference on Computer supported cooperative work (CSCW '13)* (San Antonio, TX, U.S.A.). https://doi.org/10.1145/2441776.2441941.

Roche, J. and Davis, N. (2017). 'Citizen science: an emerging professional field united in truth-seeking'. *JCOM* 16 (04), R01.
https://doi.org/10.22323/2.16040601.

Saldaña, J. (2012). The coding manual for qualitative researchers. 2nd ed. Los Angeles, CA, U.S.A.: SAGE Publications Ltd.

SciDev (2013). Global science journalism report. London, U.K.: SciDev.Net.

SciStarter (2019). *SciStarter*. URL: https://scistarter.com/.

Sheehan, K. B. (2001). 'E-mail survey response rates: a review'. *Journal of Computer-Mediated Communication* 6 (2). https://doi.org/10.1111/j.1083-6101.2001.tb00117.x.

Sloan, L., Morgan, J., Burnap, P. and Williams, M. (2015). 'Who Tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data'. *PLOS ONE* 10 (3), e0115545. https://doi.org/10.1371/journal.pone.0115545.

Statista (2018). *Countries with most Twitter users 2018*. URL: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/.

Stebbins, R. A. (1996). 'Volunteering: a serious leisure perspective'. *Nonprofit and Voluntary Sector Quarterly* 25 (2), pp. 211–224. https://doi.org/10.1177/0899764096252005.

Storksdieck, M., Shirk, J. L., Cappadonna, J. L., Domroese, M., Göbel, C., Haklay, M., Miller-Rushing, A. J., Roetman, P., Sbrocchi, C. and Vohland, K. (2016). 'Associations for citizen science: regional knowledge, global collaboration'. *Citizen Science: Theory and Practice* 1 (2). https://doi.org/10.5334/cstp.55.

Strasser, B. J., Baudry, J., Mahr, D., Sanchez, G. and Tancoigne, E. (2019). '"Citizen science"? Rethinking science and public participation'. *Science & Technology Studies* 32 (2), pp. 52–76. https://doi.org/10.23987/sts.60425.

Taylor, V. and Whittier, N. (1992). 'Collective identity in social movement communities: lesbian feminist mobilization'. In: Frontiers in social movement theory. Ed. by A. D. Morris and C. M. Mueller. New Haven, CT, U.S.A.: Yale University Press, pp. 104–129.

Torpey, J. (9th April 2013). *Gardeners needed to join army of citizen scientists*. URL: http://westerngardeners.com/gardening-projects-for-citizen-scientists.html.

Townsend, L. and Wallace, C. (2016). *Social media research: a guide to ethics*. URL: https://www.gla.ac.uk/media/media_487729_en.pdf.

Trumbull, D. J., Bonney, R., Bascom, D. and Cabral, A. (2000). 'Thinking scientifically during participation in a citizen-science project'. *Science Education* 84 (2), pp. 265–275. https://doi.org/10.1002/(SICI)1098-237X(200003)84:2<265::AID-SCE7>3.0.CO;2-5.

Twitter, Inc. (2019). *More about restricted uses of the Twitter APIs*. Developer terms. URL: https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases.

UNESCO Institute for Statistics (UIS) (2018). *Women in science*. UIS Fact Sheet, no. 51 (June), pp. 1–4.

Verkuyten, M., de Jong, W. and Masson, C. (1995). 'The construction of ethnic categories: discourses of ethnicity in the Netherlands'. *Ethnic and Racial Studies* 18 (2), pp. 251–276. https://doi.org/10.1080/01419870.1995.9993863.

Wang, T., Brede, M., Ianni, A. and Mentzakis, E. (2017). 'Detecting and characterizing eating-disorder communities on social media'. In: *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining — WSDM '17*. Cambridge, U.K.: ACM Press, pp. 91–100. https://doi.org/10.1145/3018661.3018706.

Williams, M. L., Burnap, P. and Sloan, L. (2017). 'Towards an ethical framework for publishing Twitter data in social research: taking into account users' views, online context and algorithmic estimation'. *Sociology* 51 (6), pp. 1149–1168. https://doi.org/10.1177/0038038517708140.

Zooniverse (2018). *Zooniverse*. URL: https://www.zooniverse.org/ (visited on 19th July 2019).

Author Elise Tancoigne is a post-doctoral fellow within the project "The Rise of Citizen Science: Rethinking Public Participation in Science", funded by and ERC/SNSF Consolidator Grant (BSCGI0_157787), headed by historian of science Bruno J. Strasser at the University of Geneva. E-mail: elise.tancoigne@unige.ch.

Supplementary material Available at https://doi.org/10.22323/2.18060205.