



ARTICLE



<https://doi.org/10.1057/s41599-024-02966-x>

OPEN

Citizen scientists—practices, observations, and experience

Michael O’Grady¹  & Eleni Mangina¹ 

Citizen science has been studied intensively in recent years. Nonetheless, the voice of citizen scientists is often lost despite their altruistic and indispensable role. To remedy this deficiency, a survey on the overall experiences of citizen scientists was undertaken. Dimensions investigated include activities, open science concepts, and data practices. However, the study prioritizes knowledge and practices of data and data management. When a broad understanding of data is lacking, the ability to make informed decisions about consent and data sharing, for example, is compromised. Furthermore, the potential and impact of individual endeavors and collaborative projects are reduced. Findings indicate that understanding of data management principles is limited. Furthermore, an unawareness of common data and open science concepts was observed. It is concluded that appropriate training and a raised awareness of Responsible Research and Innovation concepts would benefit individual citizen scientists, their projects, and society.

¹School of Computer Science, University College Dublin, Dublin, Ireland. ✉email: michael.j.ograde@ucd.ie; eleni.mangina@ucd.ie

Introduction

Citizen Science (CS) is increasingly viewed as a viable methodology for scientific research, either as a bottom-up initiative or as a collaboration with the professional scientific community, NGOs, or government organizations. Its importance is acknowledged in legislative contexts, for example, in the EU Open Science policy (European Commission, 2019) and the Crowdsourcing and Citizen Science Act in the USA (US Government, 2017). The importance of CS throughout history is undisputed—many famous scientists depended on alternative sources of income. The era of professional science is very much a modern phenomenon. Traditionally, CS was often perceived as an exercise in data collection. However, citizen scientists have increasingly undertaken epistemic roles such as analysis and interpretation, with the online Zooniverse platform being an exemplary model. Thus, while CS is synonymous with pursuing orthodox scientific knowledge, it is also interesting to recall that there is a countercultural dimension to CS (McQuillan, 2014). Indeed, CS is often seen as a vehicle for democratizing science, for which an effective data stewardship process is vital (de Sherbinin et al., 2021). However, the core concept of democratization of science has been challenged, e.g., Strähle and Urban (2022).

One viable but generally unexplored objective for CS communities is collaborating with national and local government agencies to influence policy. Conversely, CS offers a tool for diverse governmental agencies to engage with local communities (Cvitanovic et al., 2018). Indeed, all three courses of democratic action—monitorial, deliberative, and participatory, differ significantly and require context-sensitive, open data platforms (Ruijter et al., 2017). Citizen scientists can usefully contribute to each model—collecting, analyzing, and interpreting data to inform evidence-based policy formation. However, the data informing such policies must be of adequate quality and quantity.

CS initiatives face many challenges. Prominent among these are trust and quality assurance of data—topics well documented by the professional science community. However, professional scientists also grapple with diverse data-related issues. Other challenges include inclusivity and polarization (Cooper et al., 2021). Unlike the professional science community, CS projects often know very little about their participants, possibly due to a reluctance to collect and steward personal data (Moczek et al., 2021). However, a thorough understanding of their demographics, experiences, and understanding of data is essential if meaningful outcomes in orthodox science and policy formation are desired. Thus, this paper reports on a survey to obtain such an understanding.

Background

Currently, there is no universally agreed definition of CS. One study identified 35 definitions (Haklay et al., 2021). Such ambiguity is problematic from a policy perspective, but a narrow definition risks excluding valid activities. A need for a standardized international definition has been highlighted by Heigl et al. (2019). For this discussion, CS is considered the pursuit of scientific knowledge undertaken or contributed to by those with no direct or indirect scientific role in their professional lives. While acknowledging efforts to rebrand citizen science as community science (Lin Hunter et al., 2023), this discussion adopts a holistic approach by not willfully excluding any initiative or participant that could be reasonably categorized under the above definition.

A cursory examination of the literature confirms the popularity of CS. Thus, there has been increased interest in both citizen science practice and the contributing volunteers. A Greek study by Galanos and Vogiatzakis (2022) is probably archetypical of critical stakeholders' attitudes toward CS, such as NGOs and

government agencies. Here, awareness of the term "*Citizen Science*" is relatively low, but familiarity with the concept is high (65%). The proliferation of definitions probably contributes to this situation. While the CS concept is viewed positively, various concerns were noted, including some concerning data quality. Motivations for participating in CS initiatives have been studied in the UK (West et al., 2021), while a literature survey identified an urgent need for participant diversity if initiatives were to maximize their impact (Pateman and West, 2023). While CS is sometimes portrayed as empowering marginalized communities, it may risk reinforcing inequality unless specific contexts are carefully considered (Lewenstein, 2022). Moczek et al. (2021) surveyed the citizen science landscape in Germany and found that the level of knowledge in projects regarding contributing volunteers was shallow. As Germany is unlikely to be an outlier, this finding has profound implications for project impact. The impact of a CS project cannot be decoupled from, amongst others, the quality of the collected data.

The potential of CS to inform evidence-based decision-making and enable primary research is compromised without practical, verifiable data collection and management practices. Thus, data management within CS projects has been studied extensively in the literature. Bowser et al. (2020) surveyed CS projects, examining the entire data lifecycle and developing recommendations concerning data access and quality. Shwe (2020) considered data management in CS through the lens of the DataONE lifecycle framework, concluding that this framework only partially fulfills CS requirements. Other researchers explored data processes within CS from the perspective of data justice, concluding that citizen scientists do not benefit as much as the professional science community and governments (Christine and Thinyane, 2021). Such a conclusion highlights a power balance concern in CS projects, demanding that additional ethical decisions relating to open data practices, including data governance, be implemented (Cooper et al., 2021).

Skepticism of CS-derived data permeates the professional scientific community, compromising the potential of CS for sustainable impact. For example, the potential of CS as a complementary but non-traditional approach to helping measure the Sustainable Development Goals (SDGs) has been acknowledged. Still, data quality is highlighted as a significant obstacle (Fritz et al., 2019). While data quality may be viewed exclusively through the lens of methodology, relatively minor issues can also contribute. For example, a water quality assessment found that emotional attachment to a site contributed to overestimating water quality (Gunko et al., 2022).

One approach to reducing concerns about data quality in CS initiatives is through explicitly communicating data management practices (Stevenson et al., 2021). Alternatively, Downs et al. (2021) advocate the need for quality control and assurance throughout the entire data lifecycle, from project conception to its conclusion. The critical role of reviewers of curated data in ensuring trust by both the professional science and society is highlighted by Gilfedder et al. (2019). In the view of Balázs et al. (2021), data quality is a methodological question that is particularly challenging due to the ambiguity of the term. This debate has been ongoing for several years; see, e.g., Cruickshank et al. (2019), Ratnieks et al. (2016), Lewandowski and Specht (2015), and Bird et al. (2014).

It is increasingly acknowledged that data quality is multifaceted, and the idea that data collected by professional scientists represent the gold standard is no longer tenable. Indeed, it is argued by Binley and Bennett (2023) that there is a double standard in operation in how professional scientists view data collected by citizen scientists as biases and limitations exist in all

datasets. Other research by Mandeville et al. (2023) suggests a complementarity between data collected by professionals and participatory scientists, especially in globally protected areas. Diverse solutions have been proposed. A permissioned blockchain network could potentially manage data ownership and provenance (Lewis et al. 2022). The use of AI on mobile Apps has been demonstrated to improve quality in a birdsong CS initiative (Jäckel et al., 2023). Nonetheless, there remains a need, especially for those driving CS initiatives, to further understand participants' expectations and aspirations regarding data, especially open data (Fox et al., 2019), and Intellectual Property Rights (IPR) (Hansen et al., 2021).

In conclusion, a study by Groom et al. (2017) identified CS data as being the most restrictive due to its licensing conditions. This means that its reuse by academia, research institutes, and government agencies is limited, thus significantly reducing its potential impact. This singular instance succinctly illustrates the need for increased awareness of data issues and the adoption of robust and transparent data management policies in future CS initiatives.

Contribution. One of the earliest surveys of CS regarding data quality and approaches to validation was that of Wiggins et al. (2011), who surveyed CS projects from the Cornell Lab of Ornithology's CS email list. Contributors to this survey were mainly documented contacts for individual CS initiatives, but some were identified from an online community directory. A more recent survey was completed in 2016 by the EU Joint Research Centre (JRC). This survey adopted an online methodology and explored data management practices among citizen scientists (Schade and Tsinaraki, 2016).

The research described here both complements and differs from that described above. It is a continuation of research documented over a decade ago. It is also a response to the invitation of Schade and Tsinaraki (2016) to undertake further complementary research on data practices in CS. The scope of this survey is broader as more demographic and project-specific details are requested. However, it goes deeper into participants' understanding of data, how data is managed in their respective projects, the degree to which their projects align with the open science model (Vicente-Saez and Martinez-Fuentes, 2018), and awareness of the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles (Jacobsen et al., 2020). The survey also explores any training participants may have received in their engagement with CS initiatives.

Methodology

A survey comprising two distinct but overlapping questionnaires was designed. The survey was administered in two phases. In Phase 1, the CS community was targeted, while Phase 2 focused on the general public.

For Phase 1, a questionnaire was constructed to elicit the CS community's broad understanding and experience of CS. Questions followed seven themes—demography, their CS project, experience as a citizen scientist, data collection, data management, data dissemination, open research, including Responsible Research and Innovation (RRI), training received, etc. The design of the questionnaire was influenced by findings from the literature, including O'Grady and Mangina (2022) and Schade and Tsinaraki (2016). Structurally, the research design is descriptive, comprising a survey of 47 questions designed for quantitative analysis. Practically all questions could be answered by selecting pre-formulated answers, for example, "Yes", "No", or "I do not know". A combination of single and multiple options for answers was used. Most questions were compulsory.

An online data collection approach was implemented. The survey was constructed in Google Forms and subsequently translated into several languages—French, German, Greek, Italian, Polish, Portuguese, Spanish, and Turkish. Appropriate background information on the project and its motivations was provided, enabling potential participants to decide whether to complete the survey. It was emphasized that no identifiable or personal information would be requested or should be provided. Likewise, no identifiable details of projects were requested. Participants were informed that the data would be harnessed for scientific publications and reports. Only when participants had consented could they access the core survey. Data was only stored when the final submit button was pressed. The survey was demanding time-wise; thus, potential participants were forewarned that it would take almost 30 min to complete.

Various communication channels were harnessed to advertise the survey; these included fora for citizen scientists, including those of the European Citizen Science Association (ECSA) and Zooniverse. As the survey was anonymous, remuneration was impossible. As a token of our gratitude, a donation was made to UNICEF.

Phase 2 focused on the general public. Here, the intention is to establish a baseline for comparison. The questionnaire harnessed in Phase 1 was adopted, focusing on data concepts and training while excluding the core questions about CS. Again, the questionnaire was constructed using Google Forms. In this case, however, the services of Prolific Academic Ltd. were availed of to recruit participants. Thus, specific population characteristics could be controlled for, restricting participants geographically and ensuring gender balance. Participants were generally multilingual but listed English as a second language. The survey was relatively short, taking about six minutes on average. Again, participants were given a description of the study and how the data would be utilized and shared, and they were asked for their consent before commencing the survey.

Results

One-hundred and twenty citizen scientists completed the survey. After a rigorous quality check, 100 submissions were deemed consistent. The resultant dataset was then encoded and analyzed using MS Excel.

The gender profile constituted 53% of participants identifying as female, 45% as male, and 2% preferred not to say. The age profile ranged from 18 to 65+. The largest subgroup, 31%, was within the 35 to 44 age group. Interestingly, 13% were in the 65+ category. Participants came from 15 different European countries, with 5% from outside Europe. Over 50% of participants defined themselves as active citizen scientists, with 40% identifying as project leaders.

Biodiversity, earth science, and environmental science accounted for almost 80% of the CS projects. The geographic scope of projects ranged from neighborhood to continental, with regional (32%) and country (23%) being the most popular. Project timescales ranged from 1–4 years (44%) to more than four years (35%). Funding was generally sourced internationally (32%) and nationally (19%). Thirteen percent (13%) were unaware of how their project was funded. Academic institutions (46%) led projects, while NGOs initiated 19%. Only 55% of participants collaborated with the project leader, while 38% collaborated with people they knew only through their CS activities. Over 63% of participants contributed to project management or decision-making.

Conservation and nature protection were the primary motivations for engagement in CS (66%), followed by education and learning (62%). Most participants were active in CS for less than

five years; however, two claimed they had been active for up to 50 years. Participants contributed to all the standard CS activities, from problem definition to analysis to interpretation. Predictably, data collection was the predominant activity for 85% of participants (Fig. 1).

Mobile Apps were used by 45% of participants to collect data, while 19% used a traditional paper-based approach. Participants were well-informed as to whether their data included personal and location information. However, 20% of participants could not recall how informed consent was obtained concerning the potential use of the collected data.

Almost a quarter (26%) of participants were unaware of the existence of a data management plan. Similarly, 24% were unaware of quality control processes, while 26% were unaware that metadata or documentation was available. Notably, 43% of participants were unsure of what kind of license governed their project's data. However, 73% of participants knew there was a dedicated contact person for queries on the data collected by their projects.

In the case of data dissemination, 37% claimed that data was made publicly available as datasets, mainly in a post-processed format (34%). However, 22% of participants indicated that project data was not available to the public.

When asked about their understanding of open research, a good awareness of its principles was reported, especially concerning access and, to a lesser degree, data. Awareness of open science was surprisingly average at 54%. With the apparent exception of open innovation, participants had, in the main, encountered these terms as part of their CS activities (Fig. 2).

A good awareness of GDPR was observed. Understanding of the FAIR data principles (37%) and RRI (30%) were relatively low. As can be seen from Fig. 3, participation in CS contributed to participants' knowledge of these concepts, including GDPR.

Citizen Science Activities

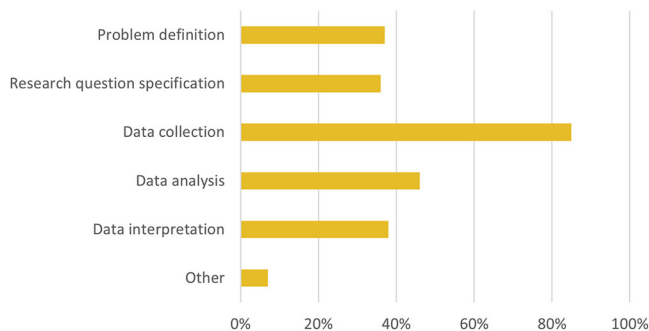


Fig. 1 How citizen scientists contribute to initiatives.

Open Research

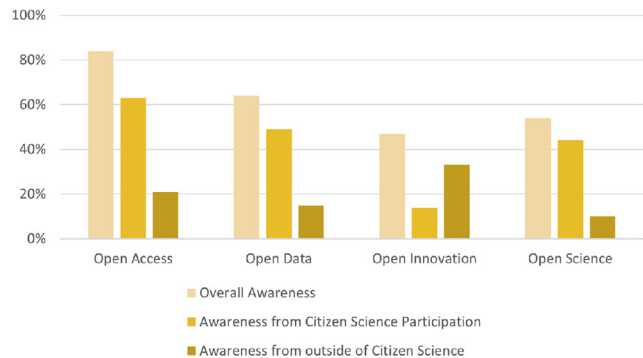


Fig. 2 Awareness of open research pillars.

To understand their experiences, participants were asked about the training they had received as part of their CS activities. Figure 4 illustrates the breakdown between formal and informal training.

Participants received training on core activities relating to CS, especially on data collection protocols, analysis, and protection. Encouragingly, training was obtained on all aspects of RRI. Apart from data collection protocols, the predominant approach to training was informal. Despite the range of training, the depth was relatively shallow in specific essential topics, including ethics, gender, and general legal issues.

A good awareness of open data repositories was demonstrated (Fig. 5). Moreover, such repositories were accessed outside CS (55%) and as part of CS initiatives (38%).

Finally, participants were asked about their views on sharing their collected data. With the notable exception of for-profit organizations, attitudes were very positive (Fig. 6).

The Public. The second phase of the survey, that of the general public, was completed by 115 participants. After the quality check, 108 participants were retained. Some corresponding results from the citizen scientists survey are included for comparative purposes. A good gender balance was observed—52% female, 47% male, with 1% preferring not to say. The age profile was dominated by the 25–34 group (38%). Participants represented 21 European countries. This survey was completed in English. All participants claimed proficiency in English, usually as a second language.

In the case of democratic models, a greater awareness of participative democracy was reported (Fig. 7), with citizen

Concepts & Terms

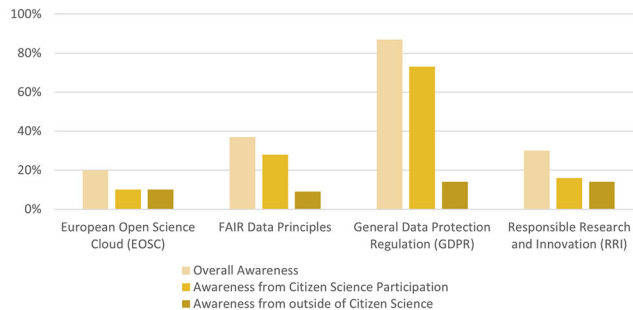


Fig. 3 Awareness of relevant concepts and terminology.

Training: Formal & Informal

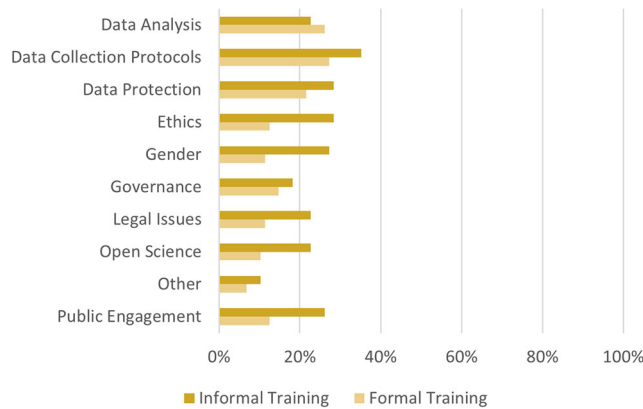


Fig. 4 Training received by participants.

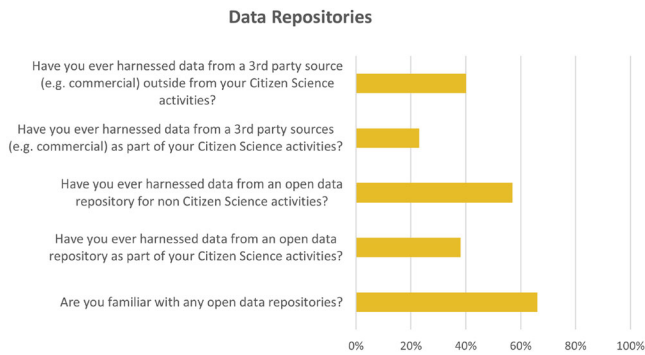


Fig. 5 Awareness and use of data repositories.

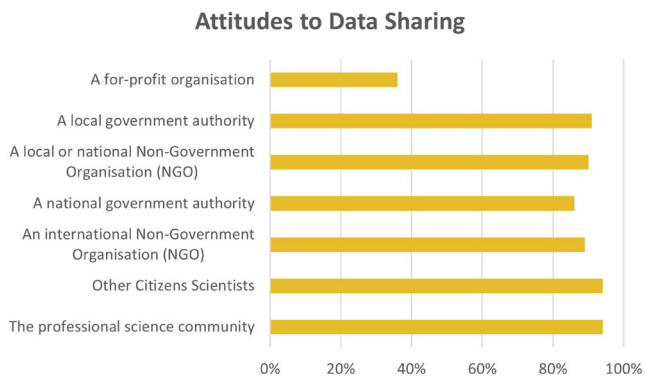


Fig. 6 Attitudes to data sharing.

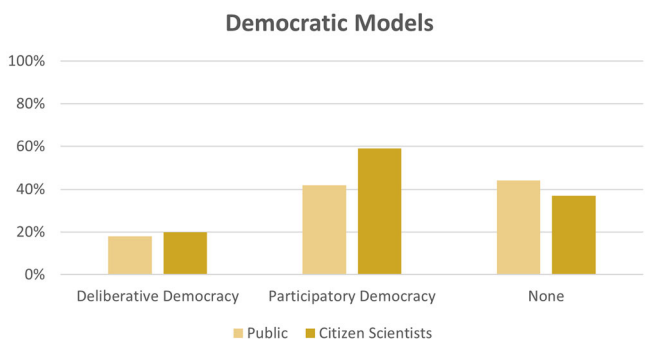


Fig. 7 Awareness of democratic models by the public and citizen scientists.

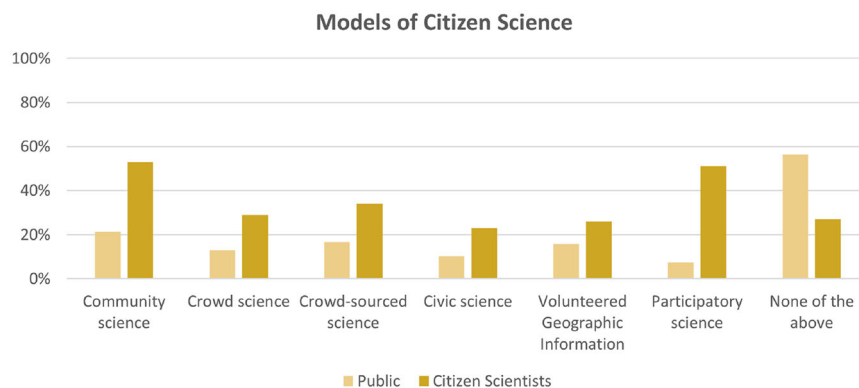


Fig. 8 Familiarity with alternative models of citizen science.

scientists being more aware of the concept (59%) than the general public (42%).

Perhaps the most striking result was that almost 88% of the European public claimed they had not encountered the term "Citizen Science". Moreover, 56% of the public stated they had not encountered any alternative models, or synonyms, of CS (Fig. 8). Over half of the citizen scientist population was familiar with the terms "community science" (53%) and "participatory science" (51%).

When offered a diverse selection of definitions of CS, the most popular was that of National Geographic for both the public and citizen scientists (Fig. 9), namely, *Citizen science is the practice of public participation and collaboration in scientific research to increase scientific knowledge. Through citizen science, people share and contribute to data monitoring and collection programs. Usually, this participation is done as an unpaid volunteer.*

Over half the general public is aware of open access (63%) and open data (56%). However, awareness of all four pillars of open research is greater amongst the CS community (Fig. 10).

The general public reported a very good awareness of GDPR (68%); however, for other concepts and terms, their knowledge was less than that of the CS population (Fig. 11).

Formal training received by the general population in data protection, ethics, and legal issues was noticeably larger than that reported by citizen scientists (Fig. 12). This pattern was replicated in the case of informal training (Fig. 13).

In each case, citizen scientists received more training in public engagement, open science, and governance—three RRI keys. In the case of gender and ethics, the general population received more training, both formal and informal.

Discussion

The implications of these results are now considered from the perspective of CS as an orthodox science methodology and a vehicle for local communities to further engage in democratic processes.

A recurring theme in the literature is the profile of the average citizen scientist—caucasian, middle-aged, college education, good socio-economic income, and male. In this survey, most participants were female and in the 35–54 age category. The results neither confirm nor challenge the stereotype of the average citizen scientist. Thus, for the credibility and integrity of arbitrary CS initiatives, inclusion remains both an objective and a priority.

CS is not a homogeneous construct. Other models, for example, community science, are broadly similar, but subtle differences and priorities may exist between them. For those seeking to harness CS, it is essential to remember that diverse communities exist. Moreover, people may not brand themselves as citizen scientists, even though their activities are archetypical of CS. It

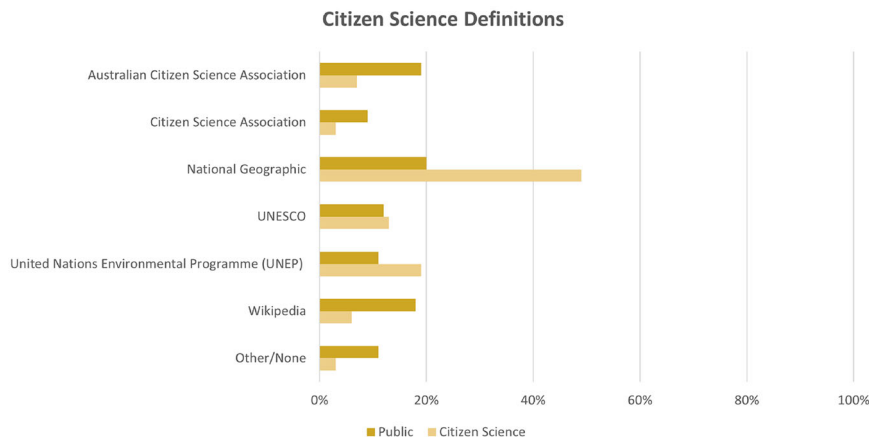


Fig. 9 Preferred definition of citizen science.

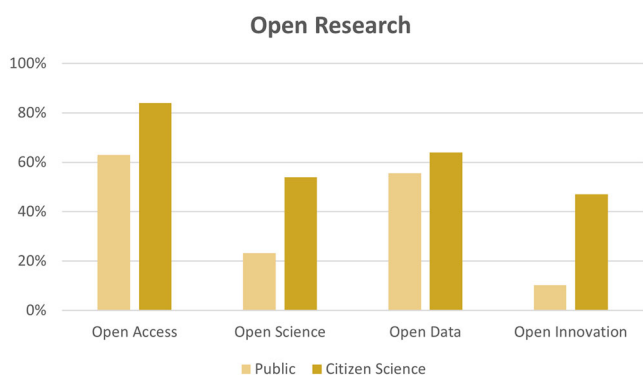


Fig. 10 Knowledge of open research pillars.

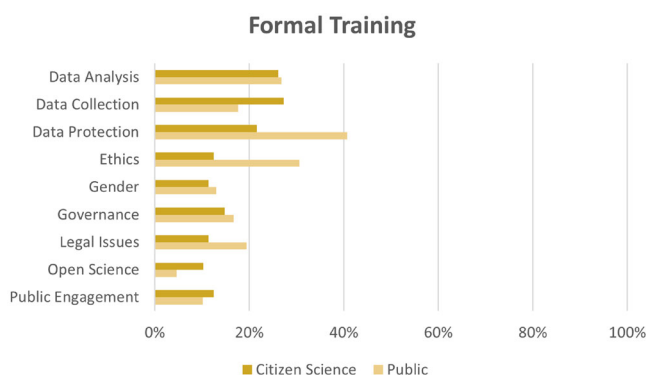


Fig. 12 Formal training received by participants.

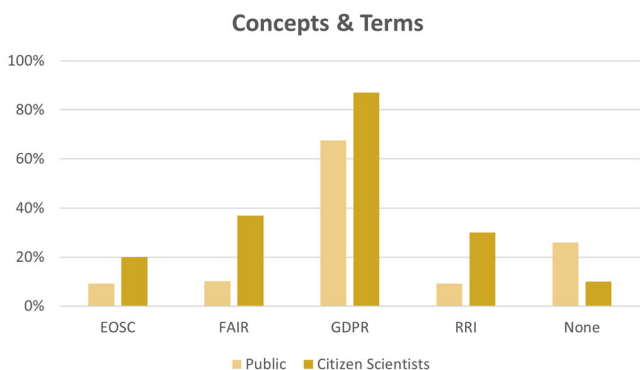


Fig. 11 Familiarity with some common open science concepts.

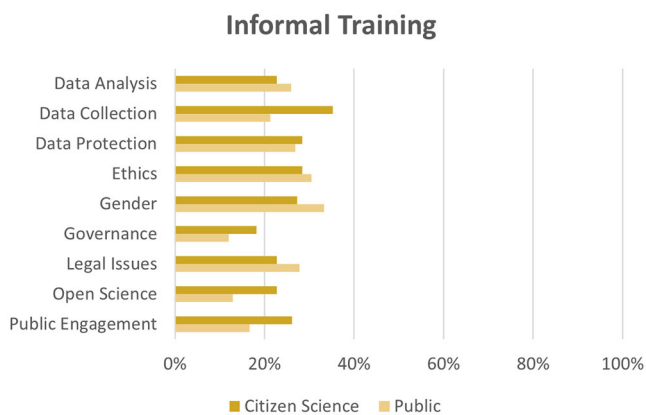


Fig. 13 Informal training received by participants.

should also be noted that awareness of these terms, as well as CS amongst the general public, seems relatively low. This would suggest that projects are poor at communicating their objectives, motivations, and activities. When promoting projects with policy objectives, an awareness of the inherent diversity within the broad CS field and its communities is essential.

While many might assume open research is essential to citizen scientists, awareness of its founding principles is mixed. Most surprisingly, "Open Science" is not especially well-known. Open Access and, to a lesser extent, Open Data are better known. However, it could be the case that these terms are understood to be synonymous with Open Science. Considering other data concepts, GDPR is well-known possibly due to the recent

emphasis on data protection. Other key concepts, such as the FAIR principles and RRI, are not well-known. However, in all cases, the general terminology is known more to the CS community than the general public.

Data collection was the most common activity by far. However, a good awareness of the sensitivity attached to identifiable and location-aware data was observed. A good understanding of how projects managed data was also observed. However, 20%-25% of participants regularly reported that they did not know or had forgotten when asked. This observation suggests that greater awareness is needed of all aspects of the data management cycle within CS initiatives. Understanding of the data licensing was

very poor, indicating that participants were unaware of how data could or could not be used going forward.

Citizen scientists received both formal and informal training as part of their preparation to engage in their respective initiatives. While the range of training received was good, the participation rate was disappointing, almost consistently less than 50%. As expected, training mainly focused on data collection and analysis protocols. However, training in other essential topics, such as ethics and open science, was limited.

Recommendations. Several recommendations have been distilled from the surveys and are applicable across initiatives, regardless of domain or motivation.

Responsible research and innovation. Many of the issues raised in this survey can be usefully considered under the broad umbrella of RRI.

- (a) **Public Engagement**—As well as the explicit goal of widening participation in projects by including diverse actors and societal stakeholders, inclusion should be interpreted broadly to include age and socio-economic profile, gender, and minorities.
- (b) **Gender**—As well as gender balance, the gender dimension should be integral to all activities.
- (c) **Education**—As well as training in diverse topics pertinent to the CS initiative under development, a more holistic approach should be followed, including the philosophy and practice of modern science. Where a policy outcome is sought, a clear understanding of how policy formation is manifested in practice should be promoted. The role of evidence, experiential and contextual, for example, in decision-making should likewise be considered.
- (d) **Open Science**—While citizen scientists are sympathetic to the objectives of Open Science, a deeper understanding of Open Research is essential to enable them to make more informed decisions and maximize the impact of their efforts.
- (e) **Ethics**—Aside from standard ethical issues covered in legislation, each project may create unique ethical issues. Additionally, ethical issues may arise during a project. It is essential that participants are aware of potential ethical issues and can recognize them as they arise. How to conduct ethical CS remains an open question (see, e.g., Rasmussen, 2021 for a treatment).
- (f) **Governance**—Inclusive of all the other RRI keys, governance remains problematic in its implementation. It should be emphasized that there is a crucial data dimension to governance (Cooper et al., 2023).

Data management. A better understanding of how data is managed within CS initiatives is needed. Training, as highlighted under RRI, is an obvious vehicle. Availability, licensing, access, sharing, and quality control should be a crucial part of any CS project briefing so that participants can make an informed decision about their contributions. Informed consent is increasingly important in the future due to the increased monetization of data (Quigley et al., 2021).

Diversity and inclusion. Citizen scientists are not a homogenous group and cannot be considered representative of an average population. Thus, inclusion is an omnipresent challenge that CS initiatives must continuously and proactively address. Crucially, any CS initiative seeking to inform policy formation must be demonstrably sensitive to the profile of its participants.

Awareness of CS. The general public lacks an understanding of CS and equivalent models. There is a need for all actors and stakeholders to promote and educate the public on the broad CS model, including its history, diverse forms, objectives, and potential. Such activities align with public engagement as considered under RRI but differ in scope and purposes.

Limitations. This study is constrained in terms of its population size. Thus, the findings cannot be regarded as definitive but rather indicative. However, the study is comparable with others in this area. For example, that of Schade and Tsinaraki (2016) attracted 121 projects. Likewise, the survey of Wiggins et al. (2011) attracted 128 project profiles but only 63 fully completed surveys. Thus, the survey reported here is typical. However, as an online survey, those who were not computer literate could not participate.

Future work. The experiences and perspectives of citizen scientists remain underexplored. This study contributed to a better understanding, but further research is needed. A complementary study but one that is deeper through replication at the country level across Europe would yield additional insights into local conditions. Such insights would provide a proper foundation for a European strategy for incorporating the CS model into policy definition and local governance.

Additional training for the CS community in diverse areas, including data management, would be beneficial. A competence framework akin to that proposed by the FabCitizen project (Pawlowski et al., 2021) could be explored further.

Finally, a deeper understanding of identity amongst citizen scientists in all their manifestations would be informative. A phenomenological study may well yield additional insights that complement the predominantly descriptive and quantitative approaches adopted to date.

Conclusion

CS is increasingly permeating the modern scientific culture. It offers intriguing possibilities to increase scientific literacy at a time when disinformation has become a widespread phenomenon. Moreover, harnessing CS and similar paradigms to inform policy and contribute to democracy is viable and intriguing. This work offers a snapshot of the experiences of citizen scientists on the ground and makes concrete recommendations as to how their contribution could be strengthened going forward. CS has had a noble history since the earliest times. With adequate support, this tradition can be continued to aid in confronting the myriad of challenges currently facing society.

Data availability

The datasets generated during this research are available from the authors upon reasonable request.

Received: 22 December 2023; Accepted: 14 March 2024;

Published online: 01 April 2024

References

- Balázs B, Mooney P, Nováková E, Bastin L, Jokar Arsanjani J (2021) Data Quality in Citizen Science. In Vohland, K. (ed.). *The science of citizen science*. Springer, Cham Switzerland. pp. 139–157
- Binley AD, Bennett JR (2023) The data double standard. *Method Ecol Evol* 14:1389–1397
- Bird TJ, Bates AE, Lefcheck JS, Hill NA, Thomson RJ, Edgar GJ, Frusher S (2014) Statistical solutions for error and bias in global citizen science datasets. *Biol Conserv* 173:144–154

- Bowser A, Cooper C, de Sherbinin A, Wiggins A, Brenton P, Chuang TR,.... Meloche M (2020) Still in need of norms: the state of the data in citizen science. *Citiz Sci Theor Pract* 5(18) <https://doi.org/10.5334/cstp.303>
- Christine DI, Thinyane M (2021) Citizen science as a data-based practice: A consideration of data justice. *Patterns* (New York, N.Y.) 2:100224
- Cooper C, Hawn C, Larson L, Parrish J, Bowser G, Cavalier D, Wilson S (2021) Inclusion in citizen science: The conundrum of rebranding. *Science* 372:1386–1388
- Cooper C, Martin V, Wilson O, Rasmussen L (2023) Equitable data governance models for the participatory sciences. *Commun Sci* 2(2). <https://doi.org/10.1029/2022CSJ000025>
- Cooper C, Rasmussen L, Jones E (2021) Perspective: the power (dynamics) of open data in citizen science. *Front Clim* 3. <https://doi.org/10.3389/fclim.2021.637037>
- Cruikshank SS, Bühler C, Schmidt BR (2019) Quantifying data quality in a citizen science monitoring program: False negatives, false positives and occupancy trends. *Conserv Sci Pract* 1(7). <https://doi.org/10.1111/csp2.54>
- Cvitanovic C, van Putten EI, Hobday AJ, Mackay M, Kelly R, McDonald J, Barnes P (2018) Building trust among marine protected area managers and community members through scientific research: Insights from the Ningaloo Marine Park, Australia. *Marine Policy* 93:195–206
- de Sherbinin A, Bowser A, Chuang T-R, Cooper C, Danielsen F, Edmunds R, Sivakumar K (2021) The critical importance of citizen science data. *Front Clim* 3:20
- Downs RR, Ramapriyan HK, Peng G, Wei Y (2021) Perspectives on citizen science data quality. *Front Clim* 3. <https://doi.org/10.3389/fclim.2021.615032>
- European Commission (2019) Open science. Open Science. Retrieved from https://research-and-innovation.ec.europa.eu/system/files/2019-12/ec_rtd_factsheet-open-science_2019.pdf
- Fox R, Bourn NAD, Dennis EB, Heafield RT, Maclean IMD, Wilson RJ (2019) Opinions of citizen scientists on open access to UK butterfly and moth occurrence data. *Biodivers Conserv* 28:3321–3341
- Fritz S, See L, Carlson T, Haklay M, Oliver JL, Fraisl D (2019) Citizen science and the United Nations sustainable development goals. *Nat Sustain* 2:922–930
- Galanos C, Vogiatzakis IN (2022) Environmental citizen science in Greece: perceptions and attitudes of key actors. *Nat Conserv* 48:31–56
- Gilfedder M, Robinson CJ, Watson JEM, Campbell TG, Sullivan BL, Possingham HP (2019) Brokering trust in citizen science. *Soc Nat Resour* 32:292–302
- Groom Q, Weatherdon L, Geijzendorffer IR (2017) Is citizen science an open science in the case of biodiversity observations? *J Appl Ecol* 54:612–617
- Gunko R, Rapeli L, Scheinin M, Vuorisalo T, Karell P (2022) How accurate is citizen science? Evaluating public assessments of coastal water quality. *Environ Policy Govern* 32:149–157
- Haklay M, Dörler D, Heigl F, Manzoni M, Hecker S, Vohland K (2021) What is citizen science? The challenges of definition. In: *Science of citizen science*. Springer, Nature, Cham. pp. 13–33
- Hansen JS, Gadegaard S, Hansen KK, Larsen AV, Møller S, Thomsen GS, Holmstrand KF (2021) Research data management challenges in citizen science projects and recommendations for library support services. A scoping review and case study. *Data Sci J* 20:25
- Heigl F, Kieslinger B, Paul KT, Uhlik J, Dörler D (2019) Opinion: toward an international definition of citizen science. *Proc Natl Acad Sci USA* 116:8089–8092
- Jäckel D, Mortega KG, Darwin S, Brockmeyer U, Sturm U, Lasseck M, Voigt-Heucke SL (2023) Community engagement and data quality: best practices and lessons learned from a citizen science project on birdsong. *J Ornithol* 164:233–244
- Jacobsen A, de Miranda Azevedo R, Juty N, Batista D, Coles S, Cornet R, Schultes E (2020) FAIR principles: interpretations and implementation considerations. *Data Intell* 2:10–29
- Lewandowski E, Specht H (2015) Influence of volunteer and project characteristics on data quality of biological surveys. *Conserv Biol* 29:713–723
- Lewenstein BV (2022) Is Citizen Science a Remedy for Inequality? *Ann Am Acad Polit Sci* 700:183–194
- Lewis R, Marstein K-E, Grytnes J-A (2022) Can blockchain technology incentivize open ecological data? <https://doi.org/10.22541/au.165425582.25694240/v1>
- Lin Hunter DE, Newman GJ, Balgopal MM (2023) What's in a name? The paradox of citizen science and community science. *Front Ecol Environ* 21:244–250
- Mandeville CP, Nilsen EB, Herfindal I, Finstad AG (2023) Participatory monitoring drives biodiversity knowledge in global protected areas. *Commun Earth Environ* 4(1). <https://doi.org/10.1038/s43247-023-00906-2>
- McQuillan D (2014) The countercultural potential of citizen science. *M/C J* 17(6). Retrieved from <https://research.gold.ac.uk/id/eprint/11482/>
- Moczek N, Hecker S, Voigt-Heucke SL (2021) The known unknowns: what citizen science projects in Germany know about their volunteers—and what they don't know. *Sustainability* 13:11553
- O'Grady M, Mangina E (2022) Adoption of responsible research and innovation in citizen observatories. *Sustainability* 14:7379
- Pateman RM, West SE (2023) Citizen science: pathways to impact and why participant diversity matters. *Citiz Sci Theor Pract* 8:50
- Pawlowski J, Nowak A, Gulevičiūtė G, Mačiulienė M (2021) O2 fabcitizen competency framework. Retrieved from <https://fabcitizen.eu/wp-content/uploads/2022/01/FabCitizen-competency-framework.pdf>
- Quigley E, Holme I, Doyle DM, Ho AK, Ambrose E, Kirkwood K, Doyle G (2021) "Data is the new oil": citizen science and informed consent in an era of researchers handling of an economically valuable resource. *Life Sci Soc Policy* 17:9
- Rasmussen LM (2021) Research ethics in citizen science. In: Iltis (ed.). *The Oxford Handbook of Research Ethics*. Oxford University Press
- Ratnieks FLW, Schrell F, Sheppard RC, Brown E, Bristow OE, Garbuzov M (2016) Data reliability in citizen science: learning curve and the effects of training method, volunteer background and experience on identification accuracy of insects visiting ivy flowers. *Method Ecol Evol* 7:1226–1235
- Ruijter E, Grimmelikhuijsen S, Meijer A (2017) Open data for democracy: developing a theoretical framework for open data use. *Gov Inf Q* 34:45–52
- Schade S, Tsinaraki C (2016) Survey report: data management in Citizen Science projects. Publication Office of the European Union, Luxembourg
- Shwe KM (2020) Study on the data management of citizen science: from the data life cycle perspective. *Data Inf Manag* 4:279–296
- Stevenson RD, Suomela T, Kim H, He Y (2021) Seven primary data types in citizen science determine data quality requirements and methods. *Front Clim* 3. <https://doi.org/10.3389/fclim.2021.645120>
- Strähle M, Urban C (2022) Why citizen science cannot answer the question of the democratisation of science. In: *Austrian Citizen Science Conference 2022*. SISSA Medialab s.r.l. Retrieved from <https://pos.sissa.it/407/001/pdf>
- US Government (2017) 15 USC 3724: Crowdsourcing and Citizen Science Act. Retrieved from <https://www.govinfo.gov/content/pkg/PLAW-114publ329/html/PLAW-114publ329.htm>
- Vicente-Saez R, Martinez-Fuentes C (2018) Open science now: a systematic literature review for an integrated definition. *J Bus Res* 88:428–436
- West S, Dyke A, Pateman R (2021) Variations in the motivations of environmental citizen scientists. *Citiz Sci Theor Pract* 6(1). <https://doi.org/10.5334/cstp.370>
- Wiggins A, Newman G, Stevenson RD, Crowston K (2011) Mechanisms for data quality and validation in citizen science. In: *The Seventh IEEE International Conference on e-Science*. CPS, IEEE, Los Alamitos, Calif. pp. 14–19

Acknowledgements

The encouragement and support of the ISEED consortium are gratefully acknowledged. The research has been supported by the European Union's Horizon 2020 research and innovation program under grant agreement No. 960366, ISEED—Inclusive Science and European Democracies (<https://iseedeurope.eu/>).

Author contributions

MO researched and defined the methodology for the survey. EM organized the collection of the data. Both authors contributed to the data analysis and the interpretation of results.

Competing interests

The authors declare no competing interests.

Ethical approval

All research was performed in accordance with relevant guidelines at University College Dublin including, but not limited to, informed consent and data protection. The research was carried out in accordance with the Declaration of Helsinki as far as applicable to this type of study. As this survey was assessed as being low risk, the exemption was obtained from the Office of Research Ethics at University College Dublin (LS-E-21-235-OGrady).

Informed consent

Informed consent was obtained from all participants.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-024-02966-x>.

Correspondence and requests for materials should be addressed to Michael O'Grady or Eleni Mangina.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024