ARPHA Proceedings

Proceedings ECSA-ÖCSK-2024, 0195-0199
doi: 10.3897/ap.e126806

Change – The transformative power of citizen science

Let's talk about data – impulses for good co-interpretation of data analysis in citizen (social) science

Sophia Segler* (a), Julia Gantenberg (a)

(a) Universität Bremen, Bremen, Germany

Abstract

This conference proceedings addresses the topic of collaborative and multi-perspective co-interpretation of data in Citizen (Social) Sciences. Co-interpretation of data is a participatory research step that is often treated with less attention in practices of citizen science and science and technology studies. However, the question emerges as to what potential entails the co-interpretation of data in Citizen Data Science. At the European Citizen Science Association (ECSA) conference in 2024, four different topics were selected for an approach to discuss best practices of co-interpreting data in Citizen Data Science. These topics cover co-interpretation of data from the perspective of different actors, necessary knowledge, data types and applicable methods and resources for collaborative data analysis in Citizen Science. The results of the four thematic groups are summarized and classified in this report.

Keywords: citizen science, citizen data science, co-interpretation, data literacy, diverse knowledge.

Published by NHM, BOKU and ECSA and peer-reviewed under responsibility of ECSA-ÖCSK-2024 (Change – The transformative power of citizen science)

Introduction

The co-interpretation of research data has the potential to not only transfer expertise and meta-knowledge between researchers and citizen scientists. We assume that knowledge is essentially situated (Haraway 1988) and that every instance of situated knowledge is shaped by socialization, social class, milieu, lifestyle, and conditioned by their respective hegemony. The incorporation of diverse knowledge as „ecology of ideas"

* Corresponding author. E-mail address: segler@uni-bremen.de

(Vickers 1968) into the interpretation of research data is contingent upon the adoption of more democratic citizen science approaches (Jaeger et al. 2023) and thus post-normal science (Funtowicz and Ravetz 1993). Furthermore, the co-interpretation of data offers a previously unexploited opportunity to transfer discipline-specific methodological knowledge and rule-based data reading skills. Consequently, data literacy is also crucial in the context of citizen science (Balázs et al. 2021; Schüller et al. 2021). In the ideal case, co-interpretation leads to greater trust in science (Kloetzer et al. 2021; Bromme 2020; Bonney 2016) due to the (self-)empowerment of citizen scientists concerning their knowledge of data handling and its scientific data interpretation. In particular, the field of social science offers a promising avenue for harnessing the potential of citizen scientists to contribute their diverse social perspectives to the interpretation of data, thereby enhancing its methodological triangulation (Flick 2011; Denzin 1970). Nevertheless, even if citizen science approaches are frequently evaluated based on their degree of participation (Eitzel et al. 2017; Haklay 2013; Shirk et al. 2012), there are still relatively few projects that also facilitate collaborative data interpretation (Heinisch 2017; Hinojosa 2021).

In the citizen social science project GINGER *(Exploring Society Together)* the "Public Data Sprint" format (Venturini et al. 2018; Segler and Gantenberg 2023) was piloted to co-interprete data on social cohesion. The objective of these intensive workshops is to interpret data in a given time, similar to hackathons in programming environments where participants collaborate to solve a problem. The objective was to disseminate this positive experience more widely in a workshop at ECSA 2024 and to discuss open questions on how to enhance the co-interpretation of data in citizen (social) sciences. In this proceeding, we will summarize the primary insights derived from the workshop discussions.

Compiling expertise for the co-interpretation of data

The workshop addressed four key topics: (1) the needs and interests of various actors involved in the co-interpretation of data (e.g., expert scientists, citizen scientists, NGOs, and other potential participants), (2) the disciplinary skills and resources required for co-interpretation of data, (3) data types that can be collaboratively analyzed, and (4) relevant methods and technologies for the co-interpretation of data.

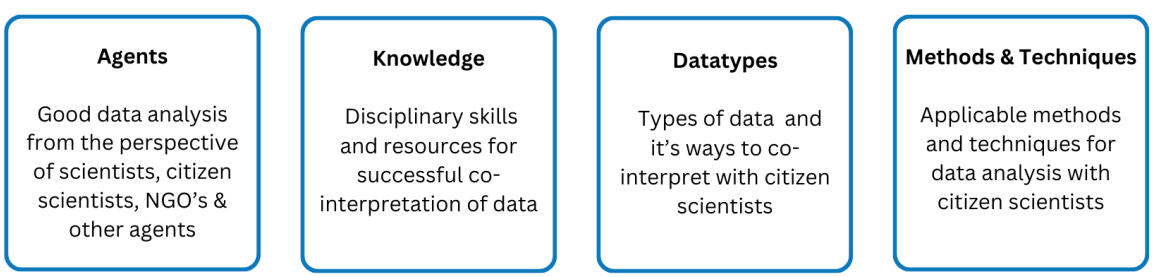

Figure 1. Workshopclusters - Agents, Knowledge, Datatypes and Methods & Techniques

Workshop-Cluster: Agents, Knowledge, Datatypes and Methods and Techniques

Upon assuming and discussing the **interests and needs of different stakeholders**, it became evident that *scientists* would benefit from the co-interpretation of data, which would lead to valuable analysis results, when applying to the FAIR principles. The Fair principles stipulate that data must be **F**indable, **A**ccessible, **I**nterpretable, and **R**eusable. Moreover, the data must also comply with ethical standards. In addition to appropriate training in data reading skills, *citizen scientists* also require an appreciative working environment in which they can work in a results-oriented manner and receive benefits and recognition for their contributions. It is assumed that *representatives of civil society organizations* will have a particular interest in analytical results that are transparent, reliable, and easy and accurate to communicate.

With regard to the **knowledge, skills, and resources required** for participatory data analysis, it was emphasized that all participants must have access to the respective knowledge to read data. The specific skills required for a given project and academic discipline will depend on the focus of the project and the discipline in question. The potential for data literacy to promote critical thinking among citizen scientists is particularly evident in the data processing and analysis phase. Consequently, it is imperative that non-experts receive training. It is of particular importance to provide accompanying materials such as guidelines, protocols, or digital training material in order to facilitate the learning process. It is therefore evident that social and pedagogical skills are of paramount importance for the effective co-interpretation with citizen scientists. Finally, tangible case studies could ensure further reflection on the optimal combination of skills and resources required for co-interpretating data.

The discussion focused on the usability of **accessible data types**, paid attention to the application of quantitative and qualitative data. In order to conduct the analysis of the previously mentioned data types, the participants of this expert group propose that co-interpretation of data can be effective when clustering the data. In particular, participation in quantitative data analysis requires technical learning environments that are easily accessible, even without programming skills, and allow for quick access. Further, it is considered important to provide stimulating moderation, professional support, and introduction to statistical programs. Moreover, the pre-selection, summarization or excerpts of data sets may be useful in order to reduce complexity. It is therefore crucial to be transparent about this procedure and to engage in critical reflection with citizen scientists.

With regard to the **methods and techniques** for effective co-interpretation of data with citizen scientists, it was proposed that the data analysis process be divided into discrete steps, a 'code of conduct' be established, and these steps be discussed in order to ensure a common understanding of the respective data analysis project. This could serve to define the most crucial conditions for scientific data analysis. Furthermore, it is decisive to consider the suitability of formats for data analysis training for non-experts, as this plays an important role in the methodological preparation for successful co-interpretation. Software recommendations for data analysis that is accessible include the open-source program JAMOVI or the social science and digital learning and analysis platform EPINetz, Orange for data mining, as well as programming environments that require in-depth specialist knowledge (at least in a supporting context) such as R, Shiny

and Python. However, it is essential that appropriate (online) tutorials are made available in order for citizen scientists to be able to deploy them at a low threshold.

In addition to numerous valuable contributions, there has been a paucity of attention paid to the potential for scientists to learn from citizen scientists when co-interpreting data in the context of peer learning. It is also important to note that no citizen scientists participated in this workshop to share their perspectives, interests, and needs on the topic of co-interpretation of data. Nevertheless, more extreme approaches of data co-interpretation can be a valuable addition to citizen science research.

Conclusion

This endeavor to engage in a more in-depth discourse on the co-interpretation of data with citizen scientists is designed to foster further resonance, to be tested and to advance knowledge based on this summary. For supplementary practical development of co-interpretation of data with citizen scientists, it is recommended that the discussion be expanded with the various specific stakeholders themselves (e. g. scientists from different disciplines, citizen scientists) in focus sessions on good data interpretation. With regard to further theoretical discussion on data co-interpretation with citizen scientists, it is evident that there is a need for analysis and discussion on more precise definitions of doing data analysis with citizens, as well as research into the processes and benefits of co-interpreting data for both citizens and vocational researchers. Finally, we intend to establish an open exchange platform, entitled “Forum Citizen Data Science”, dedicated to the discussion of topics pertaining to citizen data science in the future.

Acknowledgments

We would like to thank all workshop participants “Let’s talk about data - How can co-interpretation of data with Citizen Scientists succeed?” at ECSA 2024 as well as our student colleague Johanna Stahl, who contributed to the creation of the very creative workshop materials.

References

Balázs B, Mooney P, Nováková E, Bastin L, Jokar Arsanjani J (2021) Data Quality in Citizen Science. In: Vohland K, Land-Zandstra A, Ceccaroni L, Lemons R, Perelló J, Point M, Samson R, Wagenknecht K (Eds.) The Science of Citizen Science. Springer, Cham. https://doi.org/10.1007/978-3-030-58278-4_8

Bonney R, Phillips T B, Ballard H L, Enck J W (2016) Can citizen science enhance public understanding of science? Public Understanding of Science, 25(1), 2-16. https://doi.org/10.1177/0963662515607406

Bromme R (2022) Informiertes Vertrauen in Wissenschaft: Lehren aus der COVID-19 Pandemie für das Verständnis naturwissenschaftlicher Grundbildung (scientific literacy). Unterrichtswiss 50, 331–345. https://doi.org/10.1007/s42010-022-00159-6

Denzin N (1970) The research act. Aldine, Chicago.

Eitzel MV, Cappadonna JL, Santos-Lang C, Duerr RE, Virapongse A, West SE, Kyba CCM, Bowser A, Cooper CB, Sforzi A, Metcalfe AN, Harris ES, Thiel M, Haklay M, Ponciano L, Roche J, Ceccaroni L, Shilling FM, Dörler D, Heigl F, Kiessling T, Davis BY, Jiang Q (2017) 'Citizen Science Terminology Matters: Exploring Key Terms', Citizen Science: Theory and Practice 2(1), p. 1. https://doi.org/10.5334/cstp.96

Flick U (2011) Methoden-Triangulation in der qualitativen Forschung. VS Verlag für Sozialwissenschaften, Wiesbaden. https://doi.org/10.1007/978-3-531-92864-7_3

Funtowicz SO, Ravetz JR (1993) Science for the post-normal age. Futures, 25, 739–755. https://doi.org/10.1016/0016-3287(93)90022-L

Haklay M (2013) Citizen science and Volunteered Geographic Information: Overview and Typology of Participation. In: Sui D, Elwood S, Goodchild M (Ed.) Crowdsourcing Geographic Knowledge. Springer Netherlands, 105-122.

Haraway DJ (1988) Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. Feminist Studies, 14(3), 575–599. https://doi.org/10.2307/3178066

Heinisch B (2017) Degrees of Participation in Citizen Science Projects. An Analysis of Participatory Projects Listed in English-language and German-language Citizen Science Project Directories. In: AGES (Ed.) Austrian citizen science conference 2017. Frontiers, Wien, 15–20.

Hinojosa L, Riedy R, Polman J, Swanson R, Nuessle T, Garneau N (2021) Expanding Public Participation in Science Practices Beyond Data Collection. Citizen Science: Theory and Practice, 6(1): 11, 1–15. https://doi. org/10.5334/ cstp.292

Jaeger J, Masselot C, Greshake Tzovaras B, Senabre Hidalgo E, Haklay M, Santolini M (2023) An epistemology for democratic citizen science. Royal Society Open Science, 10 (11) , Article 231100. https://doi.org/10.1098/rsos.231100

Kloetzer L, Lorke J, Roche J, Golumbic Y, Winter S, Jõgeva A (2021) Learning in Citizen Science. In: Vohland K, Land-Zandstra A, Ceccaroni L, Lemons R, Perelló J, Point M, Samson R, Wagenknecht K (Eds.) The Science of Citizen Science. Springer, Cham. https://doi.org/10.1007/978-3-030-58278-4_15

Schüller K, Koch H, Rampelt F (2021) Data-Literacy-Charta. Version 1.1. Stifterverband, Berlin.

Segler S, Gantenberg J (2023) Innovation Data Sprint in Citizen Social Sciences. In: Heigl F, Höhener O, Dörler D (Ed.) Proceedings of the Austrian Citizen Science Conference 2023. Proceedings of Science, Sissa Medialab srl, Triest, Italien. https://doi.org/10.22323/1.442.0019

Shirk JL, Ballard HL, Wilderman CC, Phillips T, Wiggins A, Jordan R, McCallie E, Minarchek M, Lewenstein BV, Krasny ME, Bonney R (2012) Public Participation in Scientific Research: a Framework for Deliberate Design. Ecology and Society, 17(2). http://www.jstor.org/stable/26269051

Venturini T, Munk A, Meunier A (2018) Data-Sprinting: a Public Approach to Digital Research. In: Lury C, Fensham R, Heller-Nicholas A, Lammes S, Last A, Michael M, Uprichard E (Eds) Routledge Handbook of Interdisciplinary Research Methods. Routledge, London, 158 - 163. https://dx.doi.org/10.4324/9781315714523-24

Vickers G (1968) Value systems and social process. Tavistock, London.