Step Change

Mutual Learning

Focus: Data Collection

The Step Change consortium presented and discussed (with the external expert from the Barcelona Institute for Global Health representing the CitieS-Healthproject, Xavier Basagaña, and Suvodeep Mazumdar from the Sheffield University, representing the ECSA working group "Projects, Data, Tools and Technology") the key takeaways from the work on its "Citizen Science Initiatives" (CSIs) in regard to issues related to "**Data Collection**" during the project activities. Several major lessons learnt in relation to the widely challenges are listed below, along with some experiences, tested solutions and results:

In addition to the CSIs' exchange of experiences and mutual learning from this exercise, we also want Citizen Science practitioners with or without experience to be able to benefit from our findings.

The output paper was prepared by the ZSI team (Carmen Siller, ilse Marschalek, Elke Dall, Elisabeth Unterfrauner) with inputs from all consortium partners.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101006386

1. Technical and practical challenges

CSI experiences: All participating citizen science initiatives encountered some technical or practical issues during data collection. Challenges included mobilizing financial and logistical resources for deploying tools, finding appropriate locations for data collection, and changing planned data collection operations and strategies. For instance, the CSI on wildlife monitoring in Slovenia had to decide where to place camera traps, and the German CSI on energy communities had to determine which kind of data collection tool to install where. Despite being part of an EU-funded project, costs were an important issue. The question was how to achieve planned data collection within the tight budget, especially when the costs for necessary equipment turned out to be higher than originally planned due to inflation, supply chain problems, or requirements posed on the data collection tools. Project activities also need to be adapted to changing requirements, availabilities, and capabilities of citizen scientists involved in data collection – which also pose practical challenges. It was a shared experience that data collection aspects regularly need amendments, yet also that it is not easily manageable in all cases due to path dependency, and a change in data collection mechanisms in the course of the project requires a series of procedures.

Tested solutions: In the case of the German CSI, they found a feasible workaround. As the smart meter system originally selected was not the most adequate option for several households for various reasons, they modified their approach and applied an additional system, now with two parallel systems collecting data (with different costs, privacy requirements, data outputs, etc.). Although this was still not the optimal solution, they could at least continue the implementation of the CSI as generally planned and maintain the engagement of these citizen scientists. In other cases, either additional resources were found to by (extra) equipment or formal amendments were needed to the funding agency, formally changing the original project plan to respond to the changes.

Findings: CSIs have to be very flexible and adaptable to new circumstances. Strategies to implement these adaptations need to be developed, tested, and implemented. Ideally, budget allocated (and institutions supporting Citizen Science projects) also need to be flexible to help overcome these circumstances.

Key lessons learned:

- In the case of specific data collection approaches, mechanisms, or systems that do not work, a different set or a variety of tools should be developed and tested.
- As with a variety of aspects, sufficient buffer time needs to be planned to test data analysis tools.

- Data collection is a lot about people, not only about tools.
- Citizen science initiatives are dependent on funding combined with volunteer efforts, and it can be difficult to ensure their technical implementation and in particular long-term sustainability.
- It may help to have exchange with similar projects and get advice and fresh ideas in implementing technologies.

2. Methodological challenges and related skills gaps

CSI experiences: CSIs face various challenges related to finding a shared understanding of methodological and scientific terms and selecting and implementing scientific methods. To implement a CSI successfully, the implementing team needs a diverse set of skills to manage the various project cycle activities, including those related to data collection.

This includes skills related to supporting Citizen Scientists (CS) in data collection, handling data, and general methodological skills such as survey design, questionnaire development, or specific methods applied in various scientific fields. Interdisciplinary and transdisciplinary projects addressing societal challenges pose particular challenges. Participants expressed the difficulty of aligning qualitative and quantitative data collection in a methodologically appropriate way to obtain reliable data for analysis. While literature is available, many projects are not well documented, particularly with regard to the processes followed. Therefore, it is essential for new entrants in the field of CS to learn from previous experiences, which are often documented in grey literature or websites.

In addition to scientific methods, communication methods and tools are also critical. A communication campaign is necessary to inform and engage new CS, and this presents a significant challenge for CSI teams, requiring skills in information design, user experience, graphics, public relations, and campaigning. It is important to actively communicate and report about the activities to keep people engaged and involved, which is a vital but challenging task that needs to be approached in a dedicated way.

Tested solutions: The lack of expertise and skills was addressed through various means, such as consulting with more experienced peers and team members, seeking help from external experts for specific topics, and consulting with the citizen scientists themselves to find adequate solutions. Literature reviews were conducted to develop a common taxonomy, and skill gaps were addressed through training measures.

To address the challenge of aligning qualitative and quantitative data collection, the CSI took into account the different time requirements of the data collection approaches. The project was able

to combine different data collection methods, such as using automatic data inputs by sensors and asking participants to complete diaries and questionnaires (which is a time-consuming and intensive methodology of data collection). It was also helpful to apply a multi-channel approach to improve and expand communication, offering different levels of engagement[1] to the citizen scientists.

Findings: The use of physical devices for quasi-automatic data collection was found to be highly effective in engaging participants and keeping them interested. These devices measure data that is directly relevant to the citizen scientists, such as their energy consumption or the air quality in their environment. Participants expressed a strong interest in using such methods for data collection. The discussion also highlighted the availability of tools and devices on the website https://citizensciencetoolkit.eu/. Regarding taxonomies and the exchange of experiences, it was noted that relevant studies can be found using different key words, such as the "Patient Public Engagement" (PPE) terminology, which helps in accessing academic discourse on involving citizens in health-related disciplines in the UK.

The working group on "Projects, Data, Tools, and Technology"[2] of the European Citizen Science Association (ECSA) is attempting to formalize a common understanding of interests in the field. In addition to working on interoperability, data reliability, and privacy, the group also seeks to identify and share best practices and resources, exchange information about challenges, and develop standards and frameworks for citizen science data collection. The ECSA working group welcomes new members and encourages greater engagement to create further impact.

Citizen science initiatives may involve the collection of sensitive or personal data, and it is important to ensure that ethical considerations are taken into account. This may include obtaining informed consent from participants, ensuring that data is stored securely, and protecting the privacy of participants. Finally, it is important to reflect on how open and accessible citizen science data can and should be, and to emphasize the use of open science approaches and journals.

Key lessons learned:

- Training is required to instruct and support the CS teams, for example, on how to correctly use data collection devices.
- A helpdesk is needed to provide quick responses in cases of doubt, such as readily available support online or prompt replies by phone or WhatsApp, Slack, Teams etc.
- It is important to listen to the preferences of the CS for data collection and offer different tools and options accordingly.

[1] https://jcom.sissa.it/sites/default/files/figures/20/06/A02/image1.png

[2] https://www.ecsa.ngo/working-groups/projects-data-tools-and-technology/

- The different levels of engagement have to be taken into account. CS can be very active data collectors or passively collecting data – if possible, different options for data collection should be provided – with adequate strategies in place to ensure that data from different sources can be analyzed in an appropriate way, enabling the CSIs to offer conclusions and recommendations based on methodologically sound analysis of the collected data.
- In order to increase impact and support policy processes, data needs to be collected in methodologically sound ways and provided in open access.
- Standardization and frameworks are crucial and more publications and documentation about the CS process are required.

3. Challenges related to the effective use of technology and quality control

CSI experiences: Data collection often requires the use of equipment or software that citizen scientists are unfamiliar with. Sometimes, they have little to no experience with specific technologies, or, in the case of highly inclusive citizen science initiatives, with technologies such as smartphones, Apps etc. or similar. Characteristics of the target group, such as age and educational background, can also influence their familiarity with the tools. For example, some target groups may not be familiar with using smartphones or sensors.

In other cases, the selected tools may not be accessible in the language spoken by the citizen scientists or there may be barriers to installation or settings, such as on their mobile phones. These challenges and needs must be carefully addressed to ensure effective use of the technology with minimal errors. Often, not only do citizen scientists need to learn about the technology used to collect data, but those who develop and offer such tools also need to learn how to understand the requirements and needs of the users. Offering an App for example in several languages, not only the national one, could help to be more inclusive.

Based on these challenges, it is essential to control the quality of the data that is collected through the tools. In some citizen science initiatives, specific target groups have been identified as more reliable data collectors, such as those who are more experienced with the technology or with the subject of the data to be collected, for example in wildlife monitoring.

Tested solutions: Some CSIs have addressed these challenges by focusing their communication efforts on specific target groups. For example, one CSI increased outreach to younger generations, who may be more comfortable using smartphones and applications for data collection. Another approach is to dedicate resources and people to help with the installation of software, which can

be a member of the CS community or dedicated technicians. Interactive training on data collection tools, including information in local languages, has also been provided. It is important for these instructions and trainings to be concrete and focus on the concrete benefits for the CS. Workshops and meetings may need to be repeated to maintain interest over time. Providing information on how to access and analyse data, and how to use the results practically, can increase the quality and quantity of data collected and increase the sense of ownership by the CSs.

Quality control mechanisms and skills tests have been developed, such as quizzes designed to test the knowledge and skills of CSs before they begin collecting data. This can also provide additional analytical data, such as identifying difficulties in distinguishing various species in wildlife monitoring. Objective proofs, such as photos of animal sightings, can also be used for data collection, as well as comparing data entries from beginners with results from other control mechanisms, such as camera traps installed in the same region.

Debriefing sessions have also proved helpful for improving the effectiveness of data collection tools and offering further learning opportunities, such as learning more about species in the region.

Findings: Ensuring that the data collected by citizen scientists is of a high quality and meets the standards required for scientific research can be challenging. This may require the development of quality control processes and the provision of training and support to ensure that data is collected consistently and accurately.

The digital divide and inclusiveness are crucial factors to consider in each CSI. It is important to discuss different aspects of exclusion, such as who applies to participate, who is able to participate, why CS are dropping out, and what obstacles they face in participating. Some people may not be comfortable or familiar with certain technologies, which can create barriers to participation. Therefore, it is essential to remain inclusive and strike a balance between getting more and less experienced target groups involved, such as older and younger generations, experts in regional wildlife, and laypeople.

A co-design approach with test users from these groups can be used to develop data collection tools that are easy to use for everyone, including elderly people who may be motivated to participate due to environmental or health concerns. Engaging different target groups in different ways can also be effective, such as using the connections of the older generation to encourage others to participate.

Ultimately, the quality of the data collected is directly influenced by the time invested in adapting, reaching out, engaging in individual contacts, providing support, and so on.

Key lessons learned:

- The tools used for data collection should be designed in a way that promotes broad and inclusive engagement, allowing for participation from different target groups.

- To improve usability, it is important to invest time and effort in co-design approaches that take into account the needs and preferences of the target groups involved.
- Even with clear instructions provided in advance, there may still be issues that require additional support. To address this, a helpdesk that provides quick and responsive assistance is recommended.

4. Challenges related to various influences on the motivation of CS

CSI experiences: The focus of this Mutual Learning Exercise was on data collection, but in many CSIs, people are responsible for collecting data, specifically CS. As a result, the issue of recruitment reappeared. In circumstances where the involvement of CS decreases in quality or quantity, continuous engagement and motivation efforts as well as re-recruitment need to be implemented, which can be time and labor-intensive. Difficulties in recruiting and retaining specific groups or individuals, which are important for the quality of data collection efforts and to represent diverse voices, were also reported by CSIs. For instance, one CSI reported the unavailability of a specific target group to collect data. In this context, various aspects influencing the engagement of CS have been observed, such as the impact of COVID and other diseases on direct interactions with CS or contacts necessary for data collection. Furthermore, societal structures were cited as hindrances to potential participants. For example, patriarchal norms in certain communities may make it more difficult for women to engage. In some cases, socially desirable answers may be provided instead of observations.

Tested solutions: Continuity is vital to maintain motivation for activities that span a longer period of time. Acknowledging the work of citizen scientists can enhance their motivation and engagement. This can be achieved by giving them visibility, ownership, and recognition, such as highlighting a "citizen scientist of the month." Regularly thanking and acknowledging their contribution is essential, and they can also be invited to be co-authors in publications and conference papers. For instance, in the StepChange publication strategy, the possibility of involving citizen scientists in co-authorship is encouraged, and their contribution is acknowledged in compliance with authorship requirements, including discussing the order of authorship with them.

CSIs have also deployed dedicated staff to provide explanations and support to citizen scientists using various channels to answer queries, preventing frustration. If necessary, face-to-face meetings have been moved to virtual space using digital technologies or regular phone calls. Joint

meetings have also been organized in one CSI to gather support from community leaders and encourage participants, overcoming societal issues.

Findings: The balance between under-communicating and over-communicating needs to be carefully maintained, as both can have negative consequences. If requirements about data collection are unclear or open questions remain unanswered, it can lead to a lack of engagement. On the other hand, providing too much information can lead to disengagement as well. It is important to keep the specific roles and relationship with the initiative in mind. One CSI found that contacting CS 2-3 times a month with directly relevant and concrete interactions, such as trainings or dedicated meetings, was sufficient. It was also appreciated when CS could contribute to the selection of dates, and efforts were made to find suitable times for them. Tasks and activities needed to be adapted to the interests and availability of the CS. In some cases, the CS's ability to contribute exceeded expectations, especially those with prior experience who were better able to articulate their expectations and take advantage of opportunities provided.

Key lessons learned:

- Efforts need to be made to engage with hard-to-reach communities and to ensure that participants feel valued and supported.
- Support is needed in particular on technical issues to avoid frustration and keep up motivation and engagement.
- The aspect of communication cannot be underestimated.
- Formats and structures need to be flexible.