ARTICLE

<https://doi.org/10.1057/s41599-021-00851-5>

OPEN

Collaboration matters: capacity building, up-scaling, spreading, and sustainability in citizen-generated data projects

Mara Balestrini¹, Alexander Kotsev², Marisa Ponti³✉ & Sven Schade²

Projects producing citizen-generating data (CGD) to provide evidence and to drive change have increased considerably in the last decade. Many of these initiatives build on multi-actor collaboration and are often supported by non-governmental organisations (NGOs), the public sector, businesses or community-based organisations. The joint efforts of these actors are often necessary to provide the resources and the support that citizens need to produce data. In return, organisations can harness the data to support their objectives. The recent growth (or up-scaling) of CGD projects has created opportunities, as well as challenges for capacity building and sustainability. These challenges can affect the continuity and effectiveness of these initiatives and, in turn, the quality and utility of collected data. This paper analyses two CGD projects to consider their social implications and the measures necessary to increase their capacity, up-scaling, spreading, and sustainability. The case studies on noise monitoring and invasive alien species describe, respectively, a bottom-up approach at city level and a top-down approach at the European level. Regardless of the approach, capacity building requires a process of infrastructuring that engages different actors, responds to matters of concern, assesses community capacities and needs, and develops a vision and action plan. Further, the appropriation and repurposing of technical systems is required to scale up and spread CGD projects. In this process, participants' activities are shaped by technologies, while the meaning and effects of technologies are shaped through participants' activities.

¹ ESADE-Ramon Llull University, Barcelona, Spain. ² European Commission, Joint Research Centre, Ispra, Italy. ³ University of Gothenburg, Gothenburg, Sweden. ✉email: marisa.ponti@ait.gu.se

Introduction

New digital technologies, including mobile phones, sensors and online platforms, and their rapid diffusion in everyday spaces have increasingly allowed citizens to take part in projects where they collect and share data about societal issues. Such issues have included air quality, noise pollution, poor-quality housing, traffic jams, and biodiversity among many others. The subsequent growth of citizen-generated data (CGD) (DataShift, 2015) brings new opportunities for the public sector to address critical social and economic issues and improve local policy-decision making. However, significant effort is required to develop and implement CGD projects that yield useful outcomes for citizens and policy-makers. CGD projects are often multi-actor collaborations between various stakeholders, such as governments and community-based organisations, who provide the resources, support, and knowledge required for citizens to produce data.

CGD projects are similar to citizen science projects (where data is used to produce scientific knowledge and/or scientific methods are used to collect data) in terms of what they need to be “successful”. Longer term community-level involvement and activities, appropriate technological infrastructure, clear project goals and evaluation strategies, and stable funding all help to build capacity and sustain citizen science activities (Newman et al., 2012). This means that it can be difficult to scale and sustain CGD initiatives when moving from the neighbourhood to the city or even international scale. For example, it took considerable effort to re-apply air pollution project Curious Noses from the city level in Antwerp to the regional level in Flanders, Belgium (Van Brussels and Huise, 2019). Fresh Water Watch, on the contrary, is a global cyberinfrastructure that enables different kinds of local and regional projects to flourish (Hadj-Hammou et al., 2017).

Businesses are increasingly extracting value from data generated by citizens through online, mobile and Internet of Things (IoT) services in a massive, sustained and transnational manner. The large-scale use of citizen data by companies for commercial aims and profit is raising ethical concerns, which have been encapsulated by Zuboff (2019) as surveillance capitalism and considered a threat to human autonomy. There is, therefore, a pressing need to accelerate the use of CGD for the public good by understanding what factors and conditions can increase capacity building in CGD projects and improve their up-scaling, spreading, and sustainability.

This paper illustrates two CGD projects—one pan-European and one based in Barcelona, Spain—to examine how they addressed capacity building, and if they were able to scale, spread, and sustain themselves. Through the analysis of these two cases, it answers the following questions:

1. How can local CGD projects be transformed into more widespread, sustainable initiatives?
2. How can CGD projects, which normally operate at a local scale, build capacity and reach a larger cross-section of citizens at broader geographical scales?

This paper therefore contributes to the relatively understudied field of CGD by providing: (i) a description of strategies and tactics, resulting from negotiations between actors and institutions at different levels; (ii) an account of the socio-technical means through which these negotiations are conducted. It draws on Maccani and others (2020) and focuses on three interrelated constructs—infrastructuring, matters of concern, and community engagement—to describe the processes of building capacity, up-scaling, spreading, and sustainability. The two case studies allow for in-depth reflections based on first-hand experiences. The first took a bottom-up approach, built together with a local

community to address their own matter of concern, while the second took a top-down approach, planned and developed by the European Commission. Hence, the two cases complement one another, revealing commonalities as well as different challenges.

Where CGD and citizen social science intersect. This section explores the definition of CGD and its relationships with citizen science and citizen social science. CGD is a relatively new term increasingly used in the development sector, especially concerning the use of citizen monitoring approaches to achieve the Sustainable Development Goal (SDGs) (DI and DRT, 2017), which has gained acceptance by the United Nations and data communities (Haklay et al., 2020). Although the term can be used with diverse connotations in different fields (Lämmerhirt et al., 2019), this paper uses the definition of CGD as:

“Data that people or their organisations produce to directly monitor, demand or drive change on issues that affect them. It is actively given by citizens, providing direct representations of their perspectives and an alternative to datasets collected by governments or international institutions” (DataShift, 2015).

The definition provided by DataShift emphasises two main characteristics. First, the voluntary participation of the public in collecting data on issues that matter to them. CGD can be considered a form of user-generated data collected explicitly for tackling social issues, such as improving local infrastructure, tracking environmental issues, or collecting spatial data. Second, the creation of alternative datasets to complement official data, allowing citizens to make their voices heard within democratic processes at the local level (DataShift, 2015). Acting as volunteers means citizens can serve the public good (e.g., collecting data on air and water quality), and provide local authorities with evidence about problems affecting their quality of life (Ponti and Craglia, 2020). The combination of volunteer contribution and public purpose goals makes the concept of CGD similar to civic issue tracking, a participatory model in which governments are open to data generated by citizens about issues that concern them, such as street problems or disruptions due to natural disasters (Sieber and Johnson, 2015).

Given the primary focus of CGD on addressing issues of concern and producing policy outcomes, especially at the local level, the term overlaps with other forms of participation, such as community-based monitoring and policy, citizen science (Haklay et al., 2020) and citizen social science (Albert et al., 2021). The two projects included in this paper instantiate, respectively, these two forms of participation. In both cases, CGD brings new perspectives by drawing on local insights and involving communities in surfacing and responding to issues that affect them. CGD can also link to the policymaking process at the local level and contribute to redefining the interfaces between citizens, data and the public sector (Ponti and Craglia, 2020).

There are similarities and differences between CGD, citizen science, and citizen social science in terms of their orientation and focus, although the boundaries are fuzzy. Most conceptualisations of citizen science focus on the participation of citizens across different stages of a scientific project (Eitzel et al., 2017). CGP projects can include both citizen science and citizen social science, but do not necessarily have to involve participation throughout the process.

Data collected by volunteers in citizen science can also be used to formulate and implement policy to tackle complex social problems (Nascimento et al., 2018). Citizen social science can also have different purposes of knowledge production. Albert et al.

(2021) noted that citizen social sciences can focus on the participation of citizens in co-production and participatory action research. Other views focus on the engagement of citizens, local communities, and local authorities to generate data and share local knowledge, for example, for local climate policy (Kythreotis et al., 2019). Therefore, citizen social science aims to generate new scientific knowledge but also to use citizen-generated data for policymaking and governance (Albert et al., 2021).

Capacity building, up-scaling, spreading, and sustainability. Capacity building paired with political commitment is essential to support the potential of citizen science projects (Newman et al., 2012). UNDP (2009, p. 4) defines capacity building as a process “through which individuals, organisations and societies obtain, strengthen and maintain the capabilities to set and achieve their own development objectives over time.” Adapting the capacity building process recommended by the UNDP, Richter et al. (2018, p. 278) proposed the following five steps towards capacity building in citizen science: “(1) identifying and engaging different actors, (2) assessing capacities and needs for citizen science in the setting under focus, (3) developing a vision, missions and action plans, (4) developing resources such as websites and guidance, and (5) implementation and evaluation of citizen science programmes.”

Funding can be a serious challenge for capacity building in CGD projects. In citizen science, most long-term citizen science projects are supported by private and public funding through disciplines ranging from education to computer science (Crain et al., 2014). Other shared capacity building challenges include the development and maintenance of the technology necessary to support the participant community and data submission process, the recruitment of participants and the support of the infrastructure (Crain et al., 2014).

At the local scale, capacity building involves developing the conditions under which collaborating actors can learn and adapt. Local actors do not need to build something from scratch, but can leverage existing knowledge and skills to enhance and use them in new directions, driving an open and flexible process of change (Zamfir, 2017).

Three processes affected by capacity building are relevant here: sustainability, up-scaling, and spreading. Project sustainability refers to a comprehensive approach to project implementation that takes into account financial, economic, social, cultural, political, educational and environmental aspects to ensure project viability and continuity (Morfaw 2014). Maccani et al. (2020, p. 7) define project spreading as “the ability to replicate and carry over projects from one location to another at the same geographic scale - for example from one neighbourhood, city or region to another”.

Finally, project up-scaling can be defined as the potential for up-scaling and handling growing amounts of labour successfully (Bondi, 2000). CGD projects can be implemented at all scales, from local interventions to large-scale data collection efforts. The up-scaling of data production in CGD projects has been addressed in a report by the Global Partnership for Sustainable Development Data (2020), which raised the question of the quantity of data necessary to generate sufficient insights. The Global Partnership pointed out that the quantity of data needed depends on the purpose and intended user so large-scale interventions or long-term data collections are not always needed. They also noted that CGD projects scale differently according to how they organise and distribute data production (p. 29). For instance, many data collection initiatives may be combined later, further increasing the size of centrally accessible datasets. Up-scaling is challenging for many CGD projects,

Fig. 1 Conceptual framework for examining CGD projects.

although initiatives have been set up to help several projects to use new technology to generate and use data to support decision making (DataShift, 2015).

Conceptual framework. Issues related to capacity building, up-scaling, spreading, and sustainability in CGD projects have been addressed in the work conducted by DataShift, an initiative of CIVICUS, which is an international non-profit organisation. However, as noted by Meijer and Potjer (2018), their findings are neither based on academic research, nor are they positioned in the scientific literature. The field of CGD is still undertheorized and lacks understanding based on empirical evidence (Meijer and Potjer, 2018). The literature on citizen science also offers few contributions and limited insights into the factors that influence these processes (Maccani et al., 2020). To overcome these limitations, Maccani et al. (2020) drew on three theoretical approaches – diffusion of innovation (Rogers 2010), technology adoption (Venkatesh et al., 2012), and infrastructuring (Björkvinsson et al., 2010; Karasti and Baker 2008)—to identify nine conditions and characteristics, which can influence spreading and up-scaling in citizen science. This is a rare attempt to develop a theoretical framework that can be used to guide an empirical qualitative analysis of the concept of up-scaling. Therefore, it is a good point of departure to identify factors with a higher likelihood of supporting projects to scale up, spread, develop capacity-building and sustainability.

We draw upon Maccani et al. (2020) and focus on three interrelated constructs: infrastructuring, matters of concern, and community engagement. They provide a potentially useful perspective to examine the two projects as dynamic entanglements of issues, community interests, and relationships between people and technologies. All these aspects can influence the capacity building, up-scaling, spreading, and sustainability of CGD projects. Figure 1 shows these three concepts and the relationships between them. Each concept is introduced in more detail in the following sections.

Infrastructuring. Rooted in the field of Science and Technology Studies, the term infrastructuring has been used in Participatory Design to characterise the openness and adaptability required to design for uncertain outcomes and future use (Karasti, 2014; Karasti et al., 2018; LeDantec, 2012; Karasti and Baker, 2008; Star

and Bowker, 2002). In the context of the projects presented here, infrastructuring supports the growth of a project over an extended period of time, taking into account the use of existing resources (e.g., technologies) and procedures, as well as the data collected and valued by a community. Unlike the development of temporally limited projects, infrastructuring is an ongoing process of development towards the long term, where time and resources can be used in a flexible manner and diverse stakeholders can innovate together (Björgvinsson et al., 2010). LeDantec and DiSalvo (2013) used the notion of infrastructuring to sustain and advance the interests of particular communities. They suggested that this concept is useful for understanding the participatory process as something that creates or shapes a “public”. LeDantec and DiSalvo turned to Dewey in defining publics not as an a priori mass of people, but as a group of people sharing an interest in addressing a certain issue. “Constituting a public is first an expression of issues” (LeDantec and DiSalvo, 2013, p. 245) that affect and matter to that public. Thus, a public, such as a community, develops “attachments” (p. 243) to a shared issue. Understanding how a community expresses these attachments foregrounds what they think about an issue, what they need and value, what they can commit to, and how they relate to stakeholders in a specific context. The entanglement of all these aspects becomes a resource and an opportunity for engaging a public in a shared issue.

Matters of concern. The notion of a public (LeDantec and DiSalvo, 2013) in turn directs attention to the matters that concerns it. Latour (2004) distinguishes between matters of fact (what is known about phenomena) and matters of concern, which he refers to as “gatherings” (p. 233) of ideas and people around issues that matter to people. Framing participation around issues resonates with citizen social science and its orientation towards community values and local and relevant concerns (Kythreotis et al., 2019). Aligned with this view, CGD projects become techniques allowing citizens to exert their rights by producing data to evidence issues and generate better living environments (Gabrys, 2019).

Building on the notions of infrastructuring and matters of concern, Balestrini et al. (2017a) developed the Bristol Approach framework to help communities, researchers, and/or city councils plan and run sustainable and scalable CGD interventions to address local issues, including damp housing, air quality and noise pollution. The framework was developed in collaboration with stakeholders and was first used in Bristol, UK. It aims to guide stakeholders through a process of participatory infrastructuring to be enacted by community champions or facilitators, and involves co-design approaches, supporting the development of valued ownership, skills and social interactions. Building on participatory action research, Balestrini et al.’s (2017a) framework proposes six key phases: (1) identification of matters of concern, which refer to issues that a community cares and worries about; (2) framing, which involved collectively investigating the issue at stake and identifying how CGD could contribute to tackling it; (3) co-design of the project and the tools, which includes deciding at the community level what strategies should be followed, how interfaces and devices should be designed and deployed, and even what activities should take place; (4) deployment and data collection, which also includes defining a data governance protocol; (5) orchestration, meaning how to activate broader participation and raise interest about the intervention, and (6) outcomes, including how results are understood and may define future actions. The framework has already been used as the underlying methodology in a number of CGD projects to address issues ranging from odour pollution (such as Distributed Network for Odour Sensing Empowerment and Sustainability

(D-Noses) i.e., [://dnoses.eu](http://dnoses.eu)), urban environmental pollution (i.e., Making-Sense.eu) and community health (i.e., <http://citieshealth.eu>), and traffic-related pollution (<https://www.we-count.net/>).

Community engagement. An extensive review of the Human-Computer Interaction (HCI) literature on the concept of engagement led Doherty and Doherty (2018) to the conclusion that it is defined and interpreted differently by scholars. Relevant here are the characteristics of the actors, communities, and their interactions with technologies. Researchers in HCI have to manage community collaborations to address the challenge of maintaining and supporting new technologies and the difficulties that can arise after the end of community projects (Balestrini et al., 2014). HCI researchers have identified problems particularly with the up-scaling of technology interventions and the sustainability of large-scale engagement with in-situ communities (Balestrini et al., 2014). As Balestrini et al. (2017a) pointed out, reported problems include citizens’ lack of skills to use technologies and make sense of the data, and the lack of an effective approach to enthuse people around a shared purpose, developing their skills and fostering capacity building and social interactions (Balestrini et al., 2017a). To sustain community engagement, Balestrini et al. (2014) proposed a participatory approach that is collaborative and inclusive in the design, implementation, and deployment of situated HCI interventions. This approach includes five recommendations based on the results of a study investigating the factors necessary for developing sustained community ICT interventions: (1) facilitate valued ownership of projects by community stakeholders; (2) use off-the-shelf technologies in novel ways rather than novel technologies; (3) facilitate face-to-face social interactions, which can lead to discussion and ongoing engagement with the project; (4) design for participant appropriation of technologies and projects, and (5) aim for broad media coverage, which can lead to community pride and engagement. These recommendations propose interventions that are not limited to the collection of feedback from local communities but empower them to appropriate technologies in unanticipated ways, which can result in a positive sustained impact on community projects.

Methodology

The two case studies in this article benefit from the authors’ direct experiences and involvement over several years. The cases provide two distinct examples of infrastructuring CGD projects using different approaches. The bottom-up approach, the Making Sense project (Waag, 2016), is a European funded initiative deploying citizen sensing campaigns in four European cities. It emphasises the participation of the local community in the development of the project, so that they can select their own goals and the means of achieving them, starting from their matters of concern. Moreover, the community defines and plans the intervention, adapting their strategies based on change and conflict to reach a shared goal. One of the authors, Balestrini, acting as community enabler, was directly involved in the strategy and implementation of the citizen sensing and co-design activities in Barcelona. She also carried out a series of follow-up activities after the project finished.

This case study illustrates the gathering of data by citizens using low-cost sensors as a source of scientific-evidence and as part of an integrated engagement methodology. It shows how data collection together with shared analysis and interpretation can empower communities to design solutions to social challenges in a local context. It is a bottom up approach because community members self-organised. This case study presents a successful approach that also stimulated action at regional and

national levels. This demonstrates the capacity requirements when addressing the very local contexts (e.g., a single square in one city). It also shows the expectations and limitations to the spreading and up-scaling of this grassroots approach to the country level.

Unlike Making Sense, the Invasive Alien Species (IAS) uses a top-down approach combining citizen participation and system-change. IAS is a pan-European CGD project addressing a complex environmental problem and complementing collective efforts to build increased cooperation among European, national, regional, and local organisations. Two of the authors, Kotsev and Schade, are directly involved in IAS. One manages the activity in collaboration with the research centre's experts on invasive alien species and leads the technical developments of a mobile phone application and underlying infrastructure. The other was involved in the design and initial set-up of the technology system and the testing of the approach in three countries in the Danube Region. Since 2016, both have engaged regularly with stakeholders from academia, the public sector and civil society across the European Union to carry the project forward. This case study represents a leading example of a top-down (EU-institution) driven approach, in which a common issue was selected; a possible EU-wide data collection project was identified; a validation and sharing approach was proposed; and stakeholders from academia, national authorities, and civil society organisations were continuously invited to discuss and advance the proposed solution. This approach emphasises the diverse needs and opportunities across the EU, which are related to the capacities to collect and deal with the CGD available in different countries and regions.

The two cases complement each other and reveal similarities but also illustrate different challenges regarding capacity, sustainability and up-scaling.

Case studies

Each case study includes a description of the project and the role played by data and technology. It also includes the strategies used to engage communities, build capacity, and create conditions for sustainability, spread, and upscaling. An overview is presented in Table 1.

Case 1: Infrastructuring for local action. The Making Sense project aimed to make participatory sensing meaningful, sustainable and replicable to empower communities to take urban pollution into their own hands. It also aimed to co-create an open-source socio-technical toolkit for citizen-generated data on air, noise, water or radiation pollution (Woods et al., 2018). This meant making open-source technology such as sensors more accessible (affordable and useable), supporting data literacy, co-creating strategies to organise citizen sensing campaigns, and driving change (Woods et al., 2018a). Here, the focus is on a Making Sense pilot study in Plaça del Sol, Barcelona, Spain. This pilot makes a good case study because it had a specific focus and achieved tangible outcomes, bridging the gap between community engagement, data collection, policy action, and change.

The Making Sense project (2016–2017) was funded by the European Commission (Call H2020 ICT2015). It engaged more than 1000 citizens through nine experimental citizen sensing campaigns in Amsterdam, Barcelona, Maastricht and Pristina. In these campaigns, groups of citizens followed a co-creation approach in collaboration with artists, designers and scientists to develop knowledge resources and technological devices to address matters of concern.

The project was made possible through the collaboration of five European entities: Waag Society in Netherlands, the Institute of Advanced Architecture of Catalonia in Spain, and the Peer Educators Network in Kosovo, the University of Dundee in Scotland, and the EU's Joint Research Centre in Belgium. One of the key results of the project is the Making Sense toolkit, which is available online via a book with an open license (Waag, 2016).

Following the Bristol Approach framework (Balestrini et al., 2017a, 2017b), the first phase of the Making Sense project in Barcelona was to identify matters of concern in the city in collaboration with diverse communities. Three methods were used to achieve this goal: (a) a rapid ethnography process that included interviews with community organisations, residents and civic centres; (b) a review of articles in the local media reporting on environmental issues in Barcelona and a review of open city council data on citizen complaints about environmental problems; and c) three open workshops where community groups mapped environmental matters of concern in different

Table 1 Overview of the two cases.		
Name	Case study 1: Making Sense Barcelona	Case study 2: Invasive Alien Species (IAS)
Description	A European funded project to implement a replicable approach using citizen sensing to address urban environmental issues. The goal was to empower local communities affected by air, noise and nuclear pollution through open source technology and co-design methods.	A European Commission-initiated project for the early detection and monitoring of IAS (plants and animals outside their natural environment causing damage to biodiversity, agricultural production, human health, etc.). The primary aim was to support the EU Regulation on IAS.
Stakeholders	Citizens, Fab Lab Barcelona, Smart Citizen Kit community, Barcelona City Council, Making Sense partners (IAAC, Waag, University of Dundee, JRC, PEN Kosovo).	European Commission, researchers, authorities responsible for the monitoring and management of IAS (at national, regional and local levels), park managers (of natural protected areas), citizens.
Methods of Engagement	Co-design approach where participants were involved in decision making and co-creation.	Decentralised approach that acknowledges the principle of subsidiarity.
Matters of Concern	Noise pollution, air quality, and nuclear radiation.	IAS in the EU (rural areas, natural parks, cities, etc.).
Scale	Neighbourhood but scalable through replicability. Nine pilots organised at the neighbourhood level and with geographically bounded communities in Amsterdam, Maastricht, Barcelona and Pristina.	An EU-wide application that relies on country, regional and local support for its promotion and spread. Activities have been carried out in the Danube Region (including Bulgaria, Hungary, Romania, and Serbia). Dedicated support is currently being prepared for Malta and the Iberian Peninsula.
Duration	2016 to 2018, ongoing through replication by other actors.	Ongoing since 2015.
Organisational approach	Bottom-up based on a local community development.	Top-down based on the engagement of a European institution.
Link	http://making-sense.eu	https://easin.jrc.ec.europa.eu/easin/CitizenScience/BecomeACitizen

Fig. 2 Community participants in Barcelona studying how architectural features in the Plaça del Sol contribute to noise pollution and plan a sensing strategy.

neighbourhoods. In these activities, noise pollution in a public square in Gracia, a neighbourhood of Plaça del Sol, emerged as a significant matter of concern. In the final workshop of the first phase of the project, participants voted to focus on this issue.

This pilot galvanised a diverse community in Gracia, including a group of residents who claimed to have struggled with noise pollution for a decade, and a group of community champions - citizens who were trained to assist less tech-savvy participants. The pilot community comprised 35–40 neighbourhood participants, 14 community champions, 5 facilitators, 2 scientists and 2 technicians.

Following the framework, in the framing phase the community and project instigators sought to better understand the noise pollution caused by social activity and late night drinking in the square. They gathered historical documents and photographs to create a timeline of events, collecting evidence on how the square had changed and media articles demonstrating how citizens had consistently raised concerns about the use of the space (Figure 2). The participants also documented how different policies had impacted on the problem through time. Building on these assets, the community agreed that the space was now perceived exclusively as a place where young people gathered to drink at night (Waag, 2016).

In the design phase, the community developed a data collection strategy and protocol to document that noise in the Plaça del Sol was above the level established by local legislation and World Health Organisation (WHO) recommendations. Community members living around the square placed 25 Smart Citizen sensors in their flats and balconies. They co-created a sensing protocol to collect data 24 h a day for 6 weeks so they could monitor the loudest days and times. They hoped to demonstrate that noise levels were affecting their sleep and general wellbeing. Noise levels inside their homes were also recorded and sensing diaries (a paper-based diary or weekly planner) were used to record relevant information that could affect the sensor data, such as the type of windows installed, number of people in the square or waste collection schedules (Woods et al., 2018). Three experts—a community champion who was a retired sound engineer and two engineers from the Smart Citizen KIT sensor project—created a data quality strategy. The sensors were calibrated in accordance with the applicable standards using a chamber at a public university in Barcelona then deployed in participants' homes. Participants installed the sensors following clear instructions and were trained and assisted by the technical team.

Following the data collection phase (March to July 2017), three workshops and one meetup were organised to collaboratively analyse and visualise the data. The data collected by May demonstrated that noise in the Plaça averaged 70 dB around midnight. Readings tended to be even higher in the early

morning, especially Thursdays to Sundays, with peaks of more than 80 dB between 2 and 5am (Woods et al., 2018). To ensure data accuracy, the technical team supervising the analysis triangulated and compared the collected data against two official noise measuring devices managed by the city council in the area and whose data were published in a city open data portal. Moreover, the annotations taken by the participants in specific noise events were registered, and used to control for outliers and to clean the dataset.

During the action phase, empowered with the evidence that they had collected, a citizens' assembly was organised in the square. The idea was to make the data public and to invite those who use the Plaça del Sol, other neighbours and visitors to discuss the results and co-create remedial actions. Discussions were organised facilitated by experts on five themes: (i) architecture (are there any refurbishments or materials that could make the square less noisy?), (ii) urbanism and economics (how can use of the square change and be more inviting to less noisy activities? How do bars and shops in the square contribute to late night noise?), (iii) health (how is noise affecting the wellbeing of residents and what could be done about it?), (iv) children (how do children view the square and how would they like it to be?), and, finally, (v) an open table where people could spontaneously comment on the situation or offer ideas.

Around one hundred people participated and the proposed solutions were sent to the city hall, along with a report including the description of the pilot and the data analysis. Immediately afterwards, councillors and experts from the city council met with residents and a number of the co-created solutions have been implemented since then: flower benches were built where most people sat (previously preventing residents from entering their buildings and, due to the noise, from sleeping at night); an awareness and communications campaign was launched in three languages; a children's playground was installed; and the time at which the square is cleaned—making the ground wet—was moved forward, nudging people to leave at around midnight rather than 4 am.

The role of data and technology. This pilot used 25 Smart Citizen Kit sensors (Smartcitizen.me, 2021) to collect and share noise data on a daily basis. The Smart Citizen Kit is an affordable and open-source sensor device designed at Fab Lab Barcelona in 2012. It comprises an open-source Arduino board and a sensor board that can collect data on noise, air quality, temperature and humidity. The sensor connects to a database via Wi-Fi and can be deployed both indoors and outdoors. It can be powered via cord or batteries. During the first phase of the project it was observed that people with low technical skills struggled to assemble and connect the sensor. As a result, the community collaborated with

Fig. 3 Community participants in Barcelona using different techniques for data analysis.

the Smart Citizen development team to redesign the sensor setup and develop a more comprehensive onboarding interface (Balestrini et al., 2017a, 2017b). This created a strong sense of ownership over the technology and meant that participants could set up and maintain their sensors with minimal assistance.

Sensing diaries allowed participants to annotate the data following an agreed protocol. For example, if a sensor was deployed indoors, a diary note indicated if the window was open (WO) or closed (WC) at the time. These notes were key to correctly analysing the data and identifying how different activities and behaviours contributed to noise pollution.

Data was collected via the SmartCitizen.me platform where community members could choose if they wanted their sensors to show publicly or not. Devices were named by their hosts using funny nicknames and no personal data was revealed in the platform. However, in the backend each community member could access and visualise their own data, as well as those provided by fellow community members. Different tools were developed to collaboratively analyse the data, which included paper-based comparison tools and more complex data visualisation systems using Python and Jupyter Notebook (Figure 3).

Community engagement, impact, and capacity building. Three key factors contributed to fostering engagement in this project. First, the entire pilot used a co-design approach where participants were involved in decision making and co-creation from the outset, which in turn fostered their sense of agency and ownership (Balestrini et al., 2017a, 2017b). Second, pilot participants were invited to meet fortnightly at a community space in their neighbourhood to develop skills, discuss findings, take collective decisions and move the pilot forward. Face-to-face encounters are known to foster community engagement and commitment because they allow for the development of social norms and attachments. Third, at the beginning of the pilot, participants were invited to map their skills and choose how they wanted to contribute based on their availability and motivations. Tasks and resources were then distributed accordingly. Designing opportunities for people to contribute to the best of their capacity was essential in creating a truly collaborative environment where nobody felt left behind.

Most people are accustomed to using personal computers and mobile phones on a daily basis but few are familiar with sensor devices. This lack of familiarity and skills can impact how people engage with and through these technologies (DiSalvo et al., 2009). Making Sense used open source sensing technology that was made or appropriated in situ and experimental tools of this kind usually have rudimentary user interfaces. Two capacity building actions helped participants develop the skills necessary to use and maintain their sensing infrastructure. First, the use of community champions, who were trained and contributed to the co-

design of the sensing tools, allowing them to assist future participants with basic troubleshooting. Second, training sessions in which participants learnt from each other and from guests who brought knowledge and expertise.

Capacity building was instrumental in the project's impacts, of which there were several. First, soon after the pilot, the city hall deployed a communication campaign to raise awareness of how late-night noise impacted residents' sleep. This campaign consisted of street posters in Catalan, Spanish and English, as community members had identified that noise makers were often tourists who spoke foreign languages. Second, the time at which the square was cleaned changed and making the floor wet earlier reduced the number of people who congregated and sat there. Third, building on the co-created recommendations made by the participants, the city hall installed flower benches on the noisiest side of the square, which deters large numbers of people to gather there. Finally, a playground for children was installed, in hopes of attracting children back to the square and slowly changing its uses (Woods et al., 2018).

Sustainability: engagement, community orchestration, and structured funding. At the end of the campaign, participants repeatedly stated that they felt empowered. The pilot was designed from the outset with a focus on infrastructuring to support sustainability through successful community engagement and co-creation. It also followed a well-structured methodology so participants were aware of the phases and objectives; and was designed to allow ongoing onboarding, which meant that new members could join at any phase, sustaining the number of participants in case there were dropouts.

Grassroots or bottom-up approaches therefore do not mean that there is a lack of coordination or leadership (Balestrini, 2017 p. 277). Instead, bottom-up action entails the development of forms of distributed leadership and participation orchestrated in an open and horizontal way. In the Plaça del Sol Pilot, experts and volunteers coordinated the activities, engaged publics with different interests and abilities, adjusted the technologies, embedded skills in the community and facilitated social interactions. The role of facilitators and community champions is key to achieving sustained participation and engagement, as well as the effectiveness of the intervention (Balestrini, 2017).

The continued use of the pilot was assured after the completion of the project. The Making Sense project received €2M from the H2020 Programme of the European Commission. The funds were used by five institutional partners across the EU to design and implement 9 pilots. Following an infrastructuring approach, one of the aims was to produce tools and methodologies to empower grassroots groups, which are typically unfunded, through CGD. As a result, the tools developed were inexpensive, open-source and can be accessed and replicated by other technology

developers. A Toolkit was also developed and published under a Creative Commons Licence to transfer the practical knowledge needed to run a comprehensive citizen sensing campaign. Finally, a documentary demonstrating how these tools and methods were used in the context of the Making Sense pilots was openly released online. Making the technologies and methods available, and demonstrating how they can be used, has the potential to considerably reduce the costs of future citizen sensing and CGD interventions, improving their sustainability.

Horizontal spread. This openness and sharing of resources also facilitates the horizontal spread of the project. This was supported by communicating the pilot and results through word of mouth, mainstream media, and through other organisations. Other communities affected by noise pollution approached the Plaça del Sol participants seeking assistance and support. A system was organised to connect different groups and enable knowledge and technology transfer. While this was not always successful, at least two other communities in Barcelona have used the Smart Citizen Kit sensors and Making Sense methods to run CGD campaigns on noise pollution.

The documentary was screened in a local cinema and distributed online through a partnership with Playground.Media, a popular digital magazine in Spain and Latin America with an audience of over 16 million followers on Twitter. The toolkit was distributed through the online channels of the project partners and a network of interested stakeholders, including universities, living labs, and makerspaces. It was awarded an Ars Electronica STARTS Prize (2018) honourable mention and adopted at schools and institutions in Europe, Latin and North America.

Challenges. The Making Sense Barcelona pilot was successful in terms of community engagement and positive change. Nevertheless, not everything unfolded as planned; as in most participatory projects, tensions arose during the pilot. Three key challenges arose: disagreement on how to address matters of concern; differences between official datasets and CGD; and collective data governance and personal privacy.

First, the fact that community members are galvanised by a shared matter of concern does not mean they agree on the solution. In a data analysis workshop, tensions arose when participants discussed their findings. Some thought that more drastic policing measures should be demanded from the authorities, while others believed that a more subtle approach nudging people to be quiet in the square would be more socially acceptable. Such differences created discomfort and, unresolved, could have impacted the sustainability of the project. As a result, it was agreed that the community would organise an open assembly at the square and invite people to co-create potential solutions. All 17 resulting proposals were submitted to the city council for further assessment and implementation. This solution was welcomed as it meant that the final solutions blended the different perspectives.

Second, noise is measured in decibels, a unit of measurement used in acoustics and other fields to reflect its power and intensity. However, different scales exist – dBA and dBC—which have different sensitivities to various frequencies. Participants realised that their sensors recorded noise using a different scale (dBC) from the one used by the City Council in their official datasets (dBA). This concerned participants, as they wanted the data to serve as evidence of pollution, which is officially measured using dBA. Luckily, it was discovered that the City Council had one official noise sensor in the square and another nearby. The data from those sensors, recorded in dBA, was publicly available through data platform Sentilo (2021). The community integrated data from official sensors in the Smart Citizen platform to

facilitate comparisons, which, in both cases, showed that noise in the square exceeded the levels established by the WHO. After the pilot, newer versions of the Smart Citizen Kits incorporated sensors that measure in dBA.

Finally, there were tensions associated with the governance of the data. Initially, community members agreed that all data would be made publicly available online to show transparency and improve public awareness. However, as they became more familiar with the data visualisation platform, participants realised how much of their personal behaviour was revealed through the noise data. For example, children at a household discovered that their mother was louder than they were and used the data as evidence when she would tell them off for being noisy; and a couple became anxious that the sensor sitting on their windowsill would reveal their most intimate moments. The trade-off between the collective benefit of open data collided with the right to personal privacy, which triggered discussions about how to achieve the former without risking the latter. Luckily, by the time this occurred enough data had been collected to perform the analysis (6 weeks) and obtain solid evidence of sustained noise pollution in the square. As a result, the community agreed to continue collecting data following a more flexible approach where participants were free to switch their sensors on and off as needed.

Case 2: Infrastructuring for EU-wide action

Invasive Alien Species (IAS) project description. IAS facilitates the early detection and monitoring of invasive species, which are plants and animals outside their natural environment that cause damage to biodiversity, agricultural production and/or human health. The matter of concern is the need to support the EU Regulation on IAS (Regulation (EU) No. 1143/2014). The IAS project was kick-started by the Joint Research Centre (JRC) of the European Commission with the support from the European Commission's "Research and Innovation" Directorate General, as part of the MYGEOSS, a two-year project (2015–16) aimed to develop GEOSS-based (Global Earth Observation System of Systems) smart Internet applications informing European citizens on the changes affecting their local environment. This case study illustrates a top-down approach that starts capacity building at the EU-level and then engages national and regional partners to co-create solutions that suit different contexts.

Traditional scientific networks exist on this topic but the geographic coverage and temporal urgency requires complementary efforts to detect and raise awareness about IAS. For this reason, the official European scientific database of IAS – the European Alien Species Information Network (EASIN) (Tsiamis et al., 2016)—has been extended with capabilities to engage more people. These capabilities include a dedicated mobile phone application for IAS identification and reporting (Schade et al., 2019).

So far, the mobile application has been downloaded several thousand times from both the Google Play Store and the Apple App Store, and there are a few hundred active users submitting observations from across the EU. The app, which was initiated (top-down) at EU-level, is now also reused by public authorities and projects in several countries and regions, even outside the EU. There are project partners in Bulgaria, Croatia, Germany, Hungary, Malta, Portugal, Romania, Serbia and Slovenia, and Spain. The app has been and continues to be used in many ways, including complementary measures alongside the official Joint Danube Survey (the world's largest coordinated measurement of a river basin (ICPDR Secretariat, 2019)); in support of national reporting obligations; and in combination with engagements at local level, such a BioBlitzes (events aimed at finding and

Fig. 4 The architecture of the IAS in Europe (Source: Schade et al., 2019).

identifying as many species as possible in a specific area over a short period of time).

The role of data and technology. Apart from the use of dedicated environmental monitoring sensors, the widespread use of a smartphone app created a plethora of opportunities for the project. It turned out to be technically straightforward to create CGD through the dedicated IAS location-enabled app, which also benefitted from embedded sensors and smartphone cameras. There is, therefore, great potential in such solutions, especially to cover larger geographic areas.

At the same time, combining data from multiple apps remains challenging, as it is constrained by factors such as vague or absent licensing conditions, different technological solutions, and a lack of a standardised approach for the exchange of data (Adriaens et al., 2015; Schade et al., 2017). Furthermore, even if data is provided alongside licensing and access conditions, it is unclear to what extent it is reused for purposes that go beyond the initial objectives of the “native app.” This example shows that technical and organisational choices around CGD have an important impact on the reusability of the data.

The “IAS in Europe” application comprises a bundle of tools that are capable of tapping into existing geospatial data streams from citizen science but also provides the functionality to contribute new observations using a smartphone app (Schade et al., 2019). The architecture of IAS in Europe (Figure 4) consists of: (i) a backend for the storage and exposure of data through an application programming interface (API); (ii) a smartphone app that allows citizens to report sightings of invasive species and provides visualisations of existing data together with a synthesised overview of species; and (iii) a data validation interface and viewer. The technical solution is open by design, allowing its reuse on multiple levels from the local to the European.

Engaging organisations, impact, and capacity building. The development of IAS in Europe was initiated by the EU in anticipation of growing opportunities to engage citizens in scientific activities that support European policy (Science

Communication Unit, University of the West of England, 2013). The project addressed the entire EU from the onset so a waterfall model was followed to initiate and gradually extend collaborations - reaching out from the EC to EU Member States, regions, and municipalities. The project used existing European networks responsible for managing IAS in each member state as well as researchers and other stakeholders. These connections were essential to generate early feedback on the design and development of the app. Direct feedback from intermediaries, local actors, and citizens was then received as part of the dedicated study in Bulgaria, Hungary, and Romania, carried out in 2017.

Existing expert groups were also essential when moving from the pan-European level, allowing project leaders to identify additional needs of Member State authorities, the scientific community and other stakeholders. A survey was also distributed to these networks in early 2018 to gather their matters of concern as well as questions and possible critiques. Representatives of the community were then invited to a dedicated workshop in November 2018 to define the grounds for country-specific extensions and offers to engage in outreach activities and the validation of received records.

This approach enabled close collaborate with partners across the EU and continuous improvement of the application to meet their specific requirements, e.g., by including additional species, or translating the user interface of the app in another language. The open approach means that over two years the app has been used in several different contexts. These include use by national authorities for general monitoring, and data collection for dedicated projects, such as the Joint Danube Survey and the COST Action 17122 (AlienCSI) on Invasive Alien Species.

Following a stringent validation procedure, the data on IAS collected through the app has become part of the official EU-wide database maintained by EASIN (Tsiamis et al., 2016, which is used for decision and policy making at national and European levels. Figure 5 depicts how data from such projects complements official data sources. This solution is also a lead example for the guidelines on Citizen Science and Environmental Monitoring, which the EC is currently developing as part of the EU action

Fig. 5 Data on invasive species integrated into the official EU database (EASIN). The maps show the official EASIN data about IAS in the northern part of the Iberian Peninsula (top), IAS data received from Citizen Science mobile apps (middle), and the combined view of both data sources (bottom).

plan to streamline environmental reporting (COM(2017) 312 final).

The figure makes it clear how both sources complement each other because the observations cover different areas in different intensity. Also, we see that CGD were primarily collected close to urban areas (e.g., in and around Barcelona, which is located almost at the centre of the maps).

However, not all EU Member States and their authorities were aware of the opportunities and limitations when using apps to gather environmental data. The top-down approach in this case study improved understanding of how apps can be used by the EC but also by authorities and stakeholders in Member States. Setting-up an EU-wide solution and engaging with relevant communities built new capacity to embrace species records from apps in the early warning and monitoring of IAS. The project offered a starting point and guidance on how to implement working solutions. First success cases of engaging communities, along with the Joint Danube Survey, provide valuable insights and material to also engage those in other parts of the EU. It is now essential to build on these experiences to further develop the essential capacities required to mobilise local communities and individuals.

Sustainability: engagement, community orchestration, and structured funding. The project has succeeded in establishing an engaged community of practice. The JRC oversees the EU-wide coordination of the IAS project in collaboration with the Directorate General “Environment”, while national collaborations depend on the Member States and their respective structures. In this decentralised approach, partner organisations (e.g., Member State authorities, NGOs) are responsible for building or connecting with relevant communities on the ground. There is no obligation to use the app and its uptake depends on the interests and needs of different partners and communities. However, it offers a top-down solution covering technological development and maintenance, and offers full coverage of data management.

Sustainability is ensured by embedding IAS data collection in official institutional processes, and by closely engaging the community and their needs. The partners’ requirements for institutional support ensure a mandate for future developments and maintenance. They also help build relationships of trust between the many parties involved.

The sustainability of the approach depends on institutional support from the EC in overseeing app development, maintenance and use. It thereby depends on the priority and importance given to IAS policy at EU level. At the same time, an EU-wide approach provides a cost-effective solution as it does not create resource pressures on the Member States. However, the release of the different technical components under open source

licences also empower community take up and a shift of responsibilities in the future (if desired). Building the community with the direct involvement of stakeholder representatives provides solid ground for long-term evolutions and maintenance.

Up-scaling and spreading. The approach used in this case was designed to fit pan-European needs. The challenge, therefore, was not up-scaling but down-scaling to make a general EU policy relevant for countries, regions, cities, and individual citizens. Different EU Member States have different priorities, together with their own approaches and needs. On the one hand, the overall solution needed to be designed in a way that was flexible enough to be adopted in different scenarios. On the other hand, this also requires close and dedicated dialogues with Member State authorities and stakeholders. A stepwise approach from the EU-level to local actions was essential, also under the principle of subsidiarity—which demands that actions at higher administrative levels should only be carried out if they cannot be implemented at a local level.

Operating at pan-European scale, the project benefited from established legal measures, such as the General Data Protection Regulation (GDPR)—Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, which protects natural persons with regard to the processing of personal data and the free movement of such data. The project is fully compliant with the requirements on privacy protection and the handling of personal data.

In addition, flexibility was built in through a series of organisational and technical measures. Technically, unlike similar solutions, the components developed, including the app, validation interface, and data management platform: (i) are designed so that new languages or additional species can be easily added, depending on the country of use; (ii) can be decoupled from the data so that data management may be handed over to particular countries or regions; and (iii) are not rooted in a particular ICT infrastructure and can be adapted to new technological demands. As a result, different architectural approaches can be adopted depending on countries’ interests and needs. The overall approach offers a solution for CGD collection, validation and integration into official data flows. It is up to the Member States to decide if they want to take up and promote the approach in their respective country.

Regarding organisational issues, the responsible EC colleagues and existing pan-European networks were engaged even before starting any project development. The crucial measures for “down-scaling” the approach included a 2018 survey about the potential use of the application within the Member States. The survey was distributed to the networks of responsible authorities, scientists, and other stakeholders, and the dedicated follow-up

event. This contributes to capacity building but also leaves the decision of promoting the use of the application to different countries or regions. This approach enabled close collaboration with partners across the EU, and provided dedicated solutions for the Danube Region and soon for other areas such as Malta and the Iberian Peninsula.

The “IAS in Europe” app is already helping in the detection of biological invasions that are not addressed in other data sources, such as research papers and governmental repositories. The success of the project, especially the mobilisation of participants, highly depends on the local context—e.g., the capacity to promote the app in local areas such as urban environments or national parks. This type of top-down approach relies primarily on partnerships with intermediaries that take existing solutions into local and regional contexts, and also means users can be engaged in future developments and improvements.

Challenges. The establishment of the app and surrounding processes has to a large extent met, and even exceeded, initial expectations. Nonetheless, there are three aspects that did not unfold as planned. First, the initial development work behind the app and the surrounding infrastructure was by design conceptualised as open source software to establish a community of practice, which would ensure its sustainability and technological evolution. This has largely been realised with experts on biological invasions who saw the potential in reusing the existing solution instead of investing time and resources into similar developments. However, uptake by software developers remains limited, probably because the app is specialised in terms of its focus so may not appeal to developers in other fields. This in turn has led to the need to allocate resources for the update and technological evolution of the tools. A possible way of addressing this shortcoming would be to pool technical expertise from early adopters, organised through an adequate governance model.

Second, it was assumed that it would be fairly simple to reuse the lessons learned and technological tools in other domains. While this remains straightforward from a technical point of view, the migration of already developed solutions requires considerable fine-tuning, and brings new user requirements and different stakeholders. Work still needs to be done to determine an adequate approach for efficiently reusing the results in an impactful manner.

Third, and overarching the previous two points, it took considerable effort and is still an ongoing process to roll out the app across the EU so that all countries interested in using the app are supported to accommodate their country specific needs. This not only includes translating the app into different languages, but adapting the underlying data flows to the existing systems for monitoring and managing IAS, including species lists of national or regional concerns. The overall architectural design offers the required flexibility but adapting it to the needs of individual countries requires time as well as expertise and resources that are not always immediately available.

Cross-case comparison: lessons for CGD projects

This paper has examined two cases of CGD projects to explore (1) how local CGD projects can be transformed into widespread, sustainable initiatives, and (2) how CGD projects, which normally operate at the local scale, can build capacity and reach a larger cross-section of citizens at broader geographical scales.

This section extends the analysis by comparing and contrasting the insights from the two cases to answer these questions. On the one hand, their similarities highlight essential “ingredients” for successful CGD projects. On the other hand, their differences emphasise the need for projects to be relevant at the local level.

This discussion also contextualises a proposed conceptual framework to elaborate how the complex interaction of infra-structuring, matters of concern, and technology for community engagement have fostered capacity building, up-scaling, spreading, and sustainability in each project. The cross-case comparison is summarised in Table 2.

The analysis of the cases shows that the matters of concern of CGD projects are not necessarily linked to a particular local context. Unlike the locality of the matter of concern in the Plaça del Sol pilot, in the IAS the matter of concern is embodied by a binding EU legislative Act that must be applied in its entirety across the EU. The case study demonstrates the complex cooperation and coordination involved in “downscaling” the application to the national and local level, and onboarding and engaging diverse system-level organisations. Crucially, the approach has also provided a solution for CGD collection, validation and integration into official data flows.

The two cases see show how participation can be framed around such matters of concern. In fact, the projects needed the strong engagement of communities, who had expectations that needed to be met to retain their commitment, thus ensuring the sustainability of the activities that they contributed to. The Plaça del Sol activity exemplifies the importance of engaging citizens not only in data collection but also in shared data analysis and interpretation to address their matters of concern, and to reach the desired social impact. Data collection is no longer a means to an end, but part of a more sophisticated process that starts from engagement with data collection and continues to other forms of involvement, such as co-creation and co-design. Similar to what has been reported about other CGD initiatives (Lämmerhirt et al., 2019), the Plaça del Sol project has engaged the residents and educated them in the production and analysis of data. It has also been instrumental in creating a new public space between residents and local government. In the IAS case, efforts were made to involve citizens, governments, and researchers to support integrative change rather than apply a top-down state policy on an environmental threat. This was supported by a package of open reusable technologies and an open and inclusive approach for stakeholder engagement during all stages of the project starting with the scoping, implementation and validation of the results. By doing so, the aim was to empower communities to take on more responsibility for the sustainability of the project. Following Devine-Wright (2013), this example supports the finding that a local perspective on a large environmental threat can have a positive impact on citizen engagement by enhancing their sense of place attachment. CGD projects become opportunities for engaged citizens to exert their rights by producing data to evidence issues and improve their living environments (Gabrys, 2019).

The case studies illustrate how the setup and development of CGD projects are enmeshed in both technical systems (such as sensors, networks, and software), and social systems (such as people, institutions, relationships, procedures, policies, and laws). This constitutive entanglement is central in the development of CGD projects. Both case studies underline the importance of balancing different “worlds”, for example, between technology activists who helped participants to make sense of and visualise data and citizens interested in designing sensors in the Plaça del Sol pilot; and between an EU service aiming at supporting an EU Regulation and local authorities in Member States interested in adapting a technological solution to country-specific contexts.

Regarding the role of data and technologies to engage communities, it is important to note how different stakeholders—often from the private or the public sector—use the data generated by citizens. The potential for using the CGD stored in private repositories is largely understudied. In the case of Plaça

Table 2 Cross-case comparison.

Concepts	Case study	
	1: Making Sense Barcelona	2: IAS in Europe
Matters of concern	<ul style="list-style-type: none"> Local community issue 	<ul style="list-style-type: none"> EU Regulation on IAS No. 1143/2014, related regional/national regulations Environmental protection
Main goal of the community engagement	Social Impact <ul style="list-style-type: none"> Data collected by engaged citizens to be used as evidence of noise pollution. Civic mobilisation and data visualisation to increase awareness of noise pollution. Citizens' assembly to co-create potential solutions to be adopted by the city council. 	Policy Impact <ul style="list-style-type: none"> Citizen data integrated into a database used for decision- and policy-making at both national and European levels. Understanding the use of apps to collect citizen data with partner Member States and other organisations. Insights into engaging local communities and individuals.
Use of data and technology	<ul style="list-style-type: none"> Use of open-source sensing devices Co-creation approach to redesign the sensor setup Data owned and controlled by the community 	<ul style="list-style-type: none"> Adoption of an open source and reusable technical architecture and components Data made available through an open access licence Transparent approach for data reuse, incl. in authoritative sources
Infrastructuring Capacity building	Community development <ul style="list-style-type: none"> Co-design approach involving participants in decision making and co-creation. Regular meetings of participants. Task assignment based on participant skill mapping. Use of community champions helping with technology troubleshooting. 	Community development <ul style="list-style-type: none"> Engagement of system-level organisations. Reaching out to the EC, EU Member States, regions, and municipalities. Publicly available source code and well defined processes for data licensing and management acting as incentive for building capacity.
Sustainability	Engagement strategy <ul style="list-style-type: none"> Identification of matters of concern. Strengthening community relationships. Co-creation. Structured methodology. Designing for ongoing onboarding. Community orchestration <ul style="list-style-type: none"> Designing activities to foster interaction and peer-to-peer assistance Funding structure <ul style="list-style-type: none"> Project-based funding model 	Engagement strategy <ul style="list-style-type: none"> Involvement of actors across Europe and Member States. Reuse of simple open source software tools. Community orchestration <ul style="list-style-type: none"> Member State authorities or other partners free to engage with the community and promote or modify the app within their network/country. Funding structure <ul style="list-style-type: none"> Institutional (European Commission) support in overseeing the app development, maintenance, and use. Release of the source code and CGD under an open licensing scheme.
Up-scaling	<ul style="list-style-type: none"> Not applicable, because of intensity of the engagements and context dependency. Highly local issues have to be addressed where they appear. 	<ul style="list-style-type: none"> "Down-scaling" to make EU policy relevant for countries, regions, cities, and individual citizens, including translations (currently into 11 different languages).
Spreading	<ul style="list-style-type: none"> Horizontal spreading through open communication. 	<ul style="list-style-type: none"> Flexibility of the architectural solution adaptable to the interests and needs of stakeholders at multiple levels of governance.

del Sol, data ownership and control remain with the citizens who created the data in the first place. Consequently, there is a data ecosystem in which the collaboration between different parties creates value. This data ecosystem is dynamically evolving and is only partially regulated. Hence, it provides room for innovation but also brings potential threats, including data misuse or loss, and potential violations of data rights.

Another finding concerns the tensions that arose regarding data ownership when people collect data that reveal personal behaviours (e.g., noise levels in a home or balcony) and share it openly for the common good. Addressing data ownership from the onset can have a significant impact on citizen engagement, trust, and commitment, which in turn affect sustainability. However, it is important to consider that perceptions about privacy may change as participants become more knowledgeable about technology and begin to make sense of data. Although costly, projects need to be ready to adapt their protocols in response to concerns about data ownership.

The approach to infrastructuring can vary significantly across CGD projects too. In the Plaça del Sol pilot, a bottom-up approach involved the co-design of social and technical systems. Activities in actual use (for example, interpretation and articulation by citizens of the tasks they were invited to perform) intertwined with design in use activities (such as adaptation, appropriation, tailoring, re-designing, and maintenance of components) (Björgvinsson et al., 2010). Design in use was connected to the co-design of the project, as it directed attention to participants acting as co-designers and to tools that evolve to accommodate situated skills and needs. Conversely, in the IAS case, the top-down approach to infrastructuring required the coordination of a decentralised protocol to engage national and regional partners for co-creating solutions that suit different contexts. In IAS, the top-down support of an integrated infrastructure and tool allowed bottom-up activities, and allowed us to incorporate the needs of national and regional partners that joined the initiative over time.

Four main challenges affecting capacity building, up-scaling, and sustainability can be identified from the analysis of these case studies. These challenges should be taken into account when defining a strategy to go from a few successful local projects to more widespread, long-lasting initiatives, and to increase the possibilities to reach larger cross-sections of citizens at broader geographical scales.

1. **The relationship between matters of concern, engagement and sustainability.** This is complex and should be considered from the outset in accordance with a CGD project's goals. Projects that develop around a geographically situated matter of concern (as in the case of Plaça del Sol) require affected citizens to be engaged and committed to action. However, this engagement is unlikely to sustain once the matter of concern is resolved, changes or disappears. In this context, sustainability is important only as a means to achieve impact and, in the longer term, if anything, associated with the maintenance of the social bonds that can make the community more resilient in the face of future problems. In contrast, as demonstrated by the IAS project, when a CGD project deals with a matter of concern that is geographically distributed and is articulated from the top-down, it can sustain throughout time even if and when it changes or is resolved at the local level. The project is sustainable as participation is renewed each time the initiative is deployed at a new location.
2. **Engagement.** It is important to consider that CGD projects are likely to raise expectations of a response, such as from authorities (Eleta et al., 2019). However, expectations cannot always be met, and citizens may feel deceived. Therefore, Eleta et al. (2019) argued that project goals and outcomes should be set explicitly at the outset of a project to ensure informed participation, trust, and motivation. This was addressed in the Plaça del Sol project by using co-design to involve all actors in the process of goal setting and framing.
3. **Scalability.** Both cases demonstrate the importance of following a flexible approach to technology design. Open source systems and off the shelf technologies are more likely to be adopted by participants, stakeholders and collaborators, which positively impacts uptake and sustainability. However, successful adoption requires that systems are well documented and that users have the skills to successfully operate them. Moreover, the costs associated with the technology deployment and maintenance should not be underestimated. In top-down projects, such funds might be ensured by the funding entity, while this is often not the case in bottom up initiatives that can easily perish when resources are no longer available. Additionally, geographical data distribution may limit the potential for up-scaling and for homogeneous (or representative) geographic coverage of citizen sensing activities. CGD are characterised by a high degree of spatial-temporal variability, usually with data gaps in rural areas, or outside tourist areas (including, for example, natural parks). It is difficult to know if a high density of citizen observations is due to the phenomenon under study or the larger community.
4. **Emerging technologies.** These provide opportunities to generate data and respond to the results. However, without clear, widely shared licensing conditions and approaches this can result in silos of data that cannot be easily combined. Such limitations can be overcome through deliberate design choices that facilitate the reusability of the data and go beyond the objectives of a specific project. Specifically, this should involve the (i) definition and

completion of data management plans, (ii) extensive use of mainstream technology and standards, (iii) open sourcing technical developments, and (iv) adopting an open licensing scheme such as the Creative Commons.

5. **Value creation for individual citizens or communities.** Untapped potential for this remains and successful use of CGD requires a rich set of capacities including, for example, trained facilitators, thematic experts, scientists, and open-minded civil servants. As CGD projects have multiple stakeholders (Lämmerhirt et al., 2016), the limited availability of these capacities provides a major barrier when envisioning a more extensive spread or up-scaling of successful examples beyond a few localised applications.

Conclusion

This paper has focused on critical aspects for improving the capacity building, up-scaling, spreading, and sustainability of CGD projects. It has examined the strategies and tactics used in two different cases, resulting from negotiations between actors and institutions at different administrative levels. It has also provided an account of the socio-technical means through which these negotiations have been achieved. Going from a few successful local CGD projects to more widespread, long-lasting initiatives; and increasing the possibilities to reach larger cross-sections of citizens at broader geographical scales means bringing together different actors to frame matters that concern them. This process of involvement allows stakeholders to engage in a discourse around their respective points of view. Next, infra-structuring is required to develop and deploy resources such as technologies for data collection, which allow citizens to take action in support of their desired futures (LeDantec and DiSalvo, 2013).

To scale up, spread, and sustain over time, CGD projects communities need to appropriate and repurpose technical systems. Failing to engage a community around CGD at an early stage decreases the likelihood of sustainable, long-term engagement. Focusing on matters of concern is a successful approach to engaging new communities, as well as engaging existing communities. In this process, the activities of participants are shaped by technologies, while the meaning and effects of technologies are shaped through participants' activities.

This paper aims to enrich discussion about how CGD projects can be conducted and what it takes to make them more scalable, spreadable, and sustainable. The lessons learnt about community orchestration, the appropriation and use of technologies, and the need for multi-actor collaboration between different stakeholders could benefit citizen social science as well.

CGD projects are also embedded in a web (or ecosystem) of relationships and interdependencies, which results in different opportunities and constraints depending on local circumstances. The context influences the way a CGD project is valued and whether or not citizens want to participate in the first place. Exploring the opportunities and constraints involved in capacity building, up-scaling, spreading, and sustainability of CGD projects remains a topic for further research. Future work could focus on the entanglement between CGD and apps, platforms, app development environments, interface components and the associated data ecosystems on which CGD mobile apps and sensors often depend.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study. Data from the Plaça del Sol project is available from the smartcitizen.me platform (<https://smartcitizen.me/kits/tags?tags=Pla%C3%A7a>)

20del%20sol). Data from the case study on Invasive Alien Species is available from the Open Data Portal of the European Commission (<https://data.europa.eu/euodp/en/data/dataset/jrc-citsci-cs-jrc-ias>).

Received: 17 November 2019; Accepted: 24 June 2021;

Published online: 13 July 2021

References

- Albert A, Balázs B, Butkevičienė E, Mayer K, Perelló J (2021) Citizen social science: new and established approaches to participation in social research. In: Vohland K, Land-Zandstra Ceccaroni L, Lemmens R, Perelló J, Ponti M, Samson R, Wagenknecht K (eds.) *The science of citizen science*. Springer Nature, Cham, pp. 119–138
- Adriaens T, Sutton-Croft M, Owen K, Brosens D, van Valkenburg J, Kilbey D, Groom Q, Ehmig C, Thürkow F, Van Hende P, Schneider K (2015) Trying to engage the crowd in recording invasive alien species in Europe: experiences from two smartphone applications in northwest Europe. *Manag Biol Invasions* 6(2):215–225
- Balestrini M (2017) A city in common: explorations on sustained community engagement with bottom-up civic technologies. Dissertation, University College London
- Balestrini M, Rogers Y, Hassan C, Creus J, King M, Marshall P (2017a) A city in common: a framework to orchestrate large-scale citizen engagement around urban issues. In: Mark G, Fussell S, Lampe C, Schraefel MC, Hourcade JP, Appert C, Wigdor D (eds.) *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems-CHI '17*. Association for Computing Machinery (ACM), New York, pp. 2282–2294
- Balestrini M, Seiz G, Peña LL, Camprodon G (2017b). Onboarding communities to the IoT. In: Kompatsiaris I, Cave J, Satsiou A, Carle G, Passani A (eds.) *Proceedings of the 4th International Conference on Internet Science, INSCI 2017*. Springer International Publishing, Cham, pp. 19–27
- Balestrini M, Bird J, Marshall P, Zaro A, Rogers Y (2014) Understanding sustained community engagement: a case study in heritage preservation in rural Argentina. In: Jones M, Palanque P, Schmidt A, Grossman T (eds.) *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery (ACM), New York, pp. 2675–2684
- Bondi AB (2000) Characteristics of scalability and their impact on performance. In: Woodside M, Gomaa H, Menasce D (eds.) *Proceedings of the second international workshop on software and performance*. Ottawa, Canada, pp. 195–203
- Björqvinnsson E, Ehn P, Hillgren PA (2010) Participatory design and ‘democratizing innovation’. In: Robertson T, Bødker K, Bratteteig T, Loi D (eds.) *PDC '10: Proceedings of the 11th biennial participatory design conference*. Association for Computing Machinery (ACM), New York, pp. 41–50
- Crain R, Cooper C, Dickinson JL (2014) Citizen science: A tool for integrating studies of human and natural systems. *Annu Rev Environ Resour* 39 (1):641–665. <https://doi.org/10.1146/annurev-environ-030713-154609>
- DataShift (2015) What is citizen-generated data and what is the DataShift doing to promote it? http://civicus.org/images/ER%20cgd_brief.pdf. Accessed 29 Jun 2020
- Distributed Network for Odour Sensing Empowerment and Sustainability (D-Noses) (2018). <https://dnoses.eu>. Accessed 7 Oct 2020
- DiSalvo C, Louw M, Coupland J, Steiner M (2009). Local issues, local uses: tools for robotics and sensing in community contexts. In: *Proceedings of the 7th ACM Conference on Creativity and Cognition*. Association for Computing Machinery (ACM), New York, pp. 245–254. <https://doi.org/10.1145/1640233.1640271>
- Development Initiatives (DI), Development Research and Training (DRT) (2017) Citizen-generated data and sustainable development: Evidence from case studies in Kenya and Uganda. <https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/12947/citizen-generated-data-and-sustainable-development-evidence-from-case-studies-in-kenya-and-uganda.pdf>. Accessed 10 Jun 2020
- Devine-Wright P (2013) Think global, act local? The relevance of place attachments and place identities in a climate changed world. *Glob Environ Chang* 23:61–69. <https://doi.org/10.1016/j.gloenvcha.2012.08.003>
- Doherty K, Doherty G (2018) Engagement in HCI: conception, theory and measurement. *ACM Comput Surv* 51(5):99. <https://doi.org/10.1145/3234149>. (November 2018)
- Eitzel MV, Cappadonna JL, Santos-Lang C, Duerr RE, Virapongse A, West SE et al. (2017) Citizen science terminology matters: exploring key terms. *Citiz Sci Theory Pr* 2(1):1
- Eleta I, Galdon Clavell G, Righi V, Balestrini M (2019) The promise of participation and decision-making power in citizen science. *Citiz Sci Theory Pr* 4 (1):8. <https://doi.org/10.5334/cstp.171>
- Gabrys J (2019) Data citizens: how to reinvent rights. In: Bigo D, Isin E, Ruppert E (eds.) *Data politics: worlds, subjects, rights*. Routledge Studies in International Political Sociology, London, pp. 248–266
- Global Partnership for Sustainable Development Data (2020). Choosing and engaging with citizen-generated data: a guide. <http://www.data4sdgs.org/resources/choosing-and-engaging-citizen-generated-data-guide>. Accessed 20 May 2020
- Hadj-Hammou J, Loisele S, Ophof D, Thornhill I (2017) Getting the full picture: Assessing the complimentary nature of citizen science and agency monitoring data. *PLoS ONE* 12(2). <https://doi.org/10.1371/journal.pone.0188507>
- Haklay M, Fraisl D, Greshake Tzovaras B et al. (2020). Contours of citizen science: a vignette study. Available via SocArXiv. <https://doi.org/10.31235/osf.io/6u2ky>
- ICPDR Secretariat (2019) Joint Danube Survey 4. <http://www.danubesurvey.org/jds4/>. Accessed 17 Oct 2019
- Karasti H (2014) Infrastructuring in participatory design. In: *Proceedings of the Participatory Design Conference, PDC 2014*. 6–10 October 2014, Windhoek, Namibia. <https://doi.org/10.1145/2661435.2661450>
- Karasti H, Baker KS (2008). Community design: growing one's own information infrastructure. In: *PDC '08: Proceedings of the 10th Anniversary Conference on Participatory Design*, pp. 217–220. <https://doi.org/10.5555/1795234.1795280>
- Karasti H, Pipek V, Bowker GC (2018) An afterword to ‘infrastructuring and collaborative design. *Computer Supported Cooperative Work (CSCW)* 27:267–289. <https://doi.org/10.1007/s10606-017-9305-x>
- Kythreotis AP, Mantyka-Pringle C, Mercer TG, Whitmarsh LE, Corner A, Paavola J, Chambers C, Miller BA, Castree N (2019) Citizen social science for more integrative and effective climate action: a science-policy perspective. *Front Environ Sci* 7(10). <https://doi.org/10.3389/fenvs.2019.00010>
- Lämmerhirt D, Gray J, Venturini T, Meunier A (2019) Advancing sustainability together? Citizen-generated data and the Sustainable Development Goals. http://www.data4sdgs.org/sites/default/files/services_files/Advancing%20Sustainability%20Together%20CGD%20Report_0.pdf. Accessed 17 May 2019
- Lämmerhirt D, Jameson S, Prasetyo E (2016). Making citizen-generated data work. http://civicus.org/thedatashift/wp-content/uploads/2017/03/Making-Citizen-Generated-Data-Work_short-report_.pdf. Accessed 25 Feb 2019
- Latour B (2004) Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry* 30:225–248
- LeDantec CA, DiSalvo C (2013) Infrastructuring and the formation of publics in participatory design. *Soc Stud Sci* 42(2):241–264
- LeDantec CA (2012) Participation and publics: supporting community engagement. In: *CHI '12, Proceedings of the 2012 ACM annual conference on human factors in computing systems*. ACM Press, New York, pp. 1351–1360
- Maccani G, Goossens M, Righi V, Creus J, Balestrini M (2020) Up-scaling citizen science-What are the factors associated with increased reach and how to lever them to achieve impact. Publications Office of the European Union, Luxembourg, p. JRC122219. <https://doi.org/10.2760/00926>
- Meijer A, Potjer S (2018) Citizen-generated open data. An explorative analysis of 25 cases. *Gov Inf Q* 35(4):613–621
- Morfaw J (2014) Fundamentals of project sustainability. PMI® Global Congress 2014—North America. Project Management Institute, Phoenix, AZ. <https://www.pmi.org/learning/library/fundamentals-project-sustainability-9369>. Accessed 17 May 2019
- Nascimento S, Rubio Iglesias JM, Owen R, Schade S, Shanley L (2018) Citizen science for policy formulation and implementation. In: Hecker S, Haklay M, Bowser A, Makuch Z, Vogel J, Bonn A (eds.) *Citizen science: innovation in open science, society and policy*. UCL Press, London, pp. 219–240
- Newman G, Wiggins A, Crall A, Graham E, Newman S, Crowston K (2012) The future of citizen science: emerging technologies and shifting paradigms. *Front Ecol Environ* 10(6):298–304
- Ponti M, Craglia M (2020) Citizen-generated data for public policy. European Commission, Ispra, p. JRC120231, https://ec.europa.eu/jrc/communities/sites/jrccties/files/jrc120231_citizen-generated_data_for_public_policy.pdf. Accessed 31 Jan 2021
- Richter A, Dörler D, Hecker S, Heigl F, Pettibone L, Serrano F, Vohland K, Bonn A (2018) Capacity building in citizen science. In: Hecker S, Haklay M, Bowser A, Makuch Z, Vogel J, Bonn A (eds.) *Citizen Science: Innovation in Open Science, Society and Policy*. UCL Press, London, pp. 269–283. <http://www.jstor.org/stable/j.ctv550cf2.26>. Accessed 7 Feb, 2021
- Rogers EM (2010) Diffusion of innovations. Simon and Schuster, New York, NY
- Science Communication Unit, University of the West of England, Bristol (2013) Science for environment policy. In-depth report: environmental citizen science. Report produced for the European Commission DG Environment, Dec 2013
- Schade S, Tsinaraki C, Roglia E (2017) Scientific data from and for the citizen. *First Monday* 22(8). <https://doi.org/10.5210/fm.v22i7.7842>

- Schade S, Kotsev A, Cardoso A-C, Tsiamis K, Gervasini E, Spinelli F, Mitton I, Sgnaolin R (2019) Aliens in Europe: an open approach to involve more people in invasive species detection. *Comput Environ Urban Syst* 78:101384. <https://doi.org/10.1016/j.compenvurbysys.2019.101384>
- SENTILO (2021). <https://www.sentilo.io/>. Accessed 20 February, 2021
- Siebert RE, Johnson PA (2015) Civic open data at a crossroads: dominant models and current challenges. *Gov Inf Q*. 32(3):308–315
- Smartcitizen.me (2021). Smart citizen kit. <https://smartcitizen.me/>. Accessed 7 February 2021
- Star SL, Bowker G (2002) How to infrastructure? In: Lievrouw LA, Livingstone SL (eds.) *The handbook of new media. Social shaping and consequences of ICTs*. Sage, London, pp. 151–162
- START Prize (2018). <https://starts-prize.aec.at/en/making-sense-citizen-sensing-toolkit/>. Accessed 7 Feb 2021
- Tsiamis K, Gervasini E, D'Amico F et al. (2016) The EASIN editorial board: quality assurance, exchange and sharing of alien species information in Europe. *Manag Biol Invasions* 7:321–328
- Van Brussel S, Huyse H (2019) Citizen science on speed? Realising the triple objective of scientific rigour, policy influence and deep citizen engagement in a large-scale citizen science project on ambient air quality in Antwerp. *J Environ Manag* 62(3):534–551
- Venkatesh V, Thong JY, Xu X (2012) Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly* 36(1):157–178
- UNDP (2009) Supporting capacity building development. The UNDP approach. United Nations Development Programme, New York
- Waag (2016) Making sense. <https://waag.org/en/project/making-sense>. Accessed 11 Nov 2019
- Woods M, Balestrini M, Bejtullahu S, Bocconi S, Boerwinkel G, Boonstra M et al (2018). Citizen sensing: a toolkit. http://making-sense.eu/publication_categories/toolkit/. Accessed 10 Oct 2019
- Woods M, Polvora A, Hemment D, Coulson S, Nascimento S, Scott M (2018a) Making sense framework and assessment of participatory strategies. <https://op.europa.eu/en/publication-detail/-/publication/826f4f84-3169-11e8-b5fe-01aa75ed71a1/language-en>. Accessed 10 Nov 2019
- Zamfir I (2017) Understanding capacity-building/ capacity development A core concept of development policy. Available via EPRS | European Parliamentary Research Service. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2017/599411/EPRS_BRI\(2017\)599411_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2017/599411/EPRS_BRI(2017)599411_EN.pdf). Accessed 31 Jan 2021
- Zuboff S (2019) *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. Profile Books, London

Acknowledgements

The authors would like to express their gratitude to all the citizens who contributed to the case studies presented in this paper by taking part in the Making Sense and IAS projects.

The Making Sense project was funded by the European Commission, within the Call H2020 ICT2015 Research and Innovation action (GA number 688620). Ponti was funded by Marianne and Marcus Wallenberg Foundation, MMW 2018-0036.

Author contributions

All authors made substantial contributions to the conception of the work, contributed to drafting and revising the manuscript, provided final approval of the version to be published, agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved, agreed to be named on the author list, and approved the full author list.

Funding

Open access funding provided by University of Gothenburg.

Competing interests

'The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission'. Some of the authors were involved in the projects described in this paper.

Additional information

Correspondence and requests for materials should be addressed to M.P.

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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021