# EUROPEAN POLICY BRIEF

**How citizen science data can help public institutions and civil society organisations make better local decisions**

## INTRODUCTION

Citizen Science is an ever increasing, large scale and diverse approach that is crucial for public engagement due to its potential to support behaviour change at scale. It also helps generate good quality novel data and to generate such data in large quantities. For example, what started in 2016 as a competition between the two cities of Los Angeles and San Francisco, has developed into an international event, City Nature Challenge, motivating people around the world to document biodiversity and wildlife in their own cities. In 2023, over the course of one weekend, just under 2 million observations were recorded of over 57k species and over 66k people took part.[1] What does this mean for local decision making? It can help to fill data gaps; it bring perspectives of hard-to-reach groups to policy issues; it can enhance the geographical granularity of datasets; and combining citizen science data with existing official datasets to cross-validate findings can drive up data quality.

This policy brief makes the case for the use of citizen science data by local decision makers in cities, municipalities, regional environment agencies and civil society organisations. It aims to foster the use of citizen science data to inform evidence-based policies and to promote dialogue between stakeholders at multiple levels to promote better data management standards and interoperability. This policy brief will be of particular relevance to decision makers responsible for collecting, using and sharing new sources of data related to a broad range of activities spanning everything from public health to environmental monitoring.

## THE BENEFITS OF CITIZEN SCIENCE DATA TO DECISION MAKING

Citizen science is an approach that involves members of the public in voluntarily contributing to research, including by asking research questions, collecting and/or analysing data, and using the results. Citizen science projects can be initiated with a range of goals and outcomes in mind. For example, in the CompAir project citizens collect air quality data across Europe using easy-to-use sensors supplied by the project.[2] This has helped to identify hotspots of poor air quality in specific neighbourhoods leading to. Citizens have celebrated many successes in using their air quality measurements to changes in local and regional policies.[3] The unique characteristics of citizen science mean it both engages people and empowers them, augmenting traditional monitoring as people become active in their local environment. Data generated by citizen science groups have become an increasingly important source for scientists, and those pursuing the 2030 Agenda for Sustainable Development. Citizen science data are used extensively in studies of biodiversity and pollution; crowdsourced data are being used by UN operational agencies; and citizen scientists are providing data relevant to monitoring the sustainable development goals (SDGs).

[1] https://www.citynaturechallenge.org/
[2] https://www.wecompair.eu/
[3] https://eurocities.eu/latest/the-power-of-citizen-science-to-tackle-the-pollution-crisis/

Using citizen science data can have multiple benefits:

- understanding an issue in a more nuanced way while also engaging and educating people
- monitoring and reporting on sustainability targets, in a context of constrained resources, and needing to do more for less;
- broadening the scope and coverage of existing datasets, adding depth, context and nuance where government data exists, and filling gaps in areas where it doesn't;
- validating or cross-validating data, or complementing existing data with larger samples; and
- improving public perceptions of government data initiatives, and helping to build public trust in governments and public institutions.

## WHAT IS CITIZEN SCIENCE DATA?

As the field of citizen science gains recognition and momentum, more data is being created across a range of topics and sectors. As a result, there are an increasing number of opportunities to use data in policy. Different types of citizen science data include:

***On-site observations*** where citizens describe sites/locations to collect new data or enhance existing information about places, physical infrastructure, environmental conditions, wildlife presence or events. Observations range from tracking wildlife via cameras for biodiversity monitoring, to documenting the status of water and sanitation infrastructure in public facilities. Example projects include many of those found on the *Bürger schaffen Wissen* (Citizens create knowledge) citizen science platform in Germany.[4]

***Sample collection and measurement*** where citizens identify and collect different biological or environmental samples such as soil, water, or air samples, that cannot directly be observed (such as radiation) or cannot otherwise be quantified (such as temperature or noise). An example is the Isala project at the University of Antwerp to map the vaginal microbiome of healthy women.[5]

***Audio-visual recording*** through active or passive sensing, where people make audio and video recordings, collected via stationary devices such as sensors and cameras, mobile devices such as drones or via people's consumer devices such as mobile phones and cameras. An example is the Urban Belonging project in Copenhagen using participatory mapping and photography to document local residents' relationship to the city.[6]

***Classifying / tagging*** usually done remotely via online interfaces, means that people classify existing data sources such as images, sounds, video and other data, to extract meaning and add information, such as in the Koster Seafloor Observatory project in Sweden.[7] Some projects, such as Humanitarian OpenStreetMap[8], where volunteers map the impacts of natural disasters, combine an easy-to-use interface, task instructions, in combination with an accreditation system for contributors, and a peer-reviewed validation system to coordinate who classifies data and who validates it.

***Compiling data*** helps to add meaning and insights to unstructured and structured data by providing a central access point, a database, or an API. Compiling is often a necessary step towards other analytical tasks that are not possible with individual datasets, such as data definition at the beginning of a project, pattern recognition, cross-verification or others. For example, citizen groups collect high resolution aerial imagery, and put it into OpenStreetMap to be able to annotate the images with digital building footprint data.

***Triangulation*** refers to cross-verifying data with other data to improve the reliability and accuracy. Governments may use citizen science data as a control value to test the accuracy of its existing data and predictive models. In some cases, citizen science data can provide comparative data and first baselines that governments later verify by conducting their own data collection. For example, data from the UK Met Office's Weather Observation Website[9] has been used to fill observational gaps on rainfall data.

***Pattern recognition*** involves citizens in potentially discovering spatial distributions of data, such where buildings with higher exposure to disasters are located in cities; or how many households can reach public services. In other cases, citizens may discover temporal distributions such as pollution spikes at certain points in time, or continuously high air pollution values. See for example, the Science in the City project in the Barbican Estate, in central London.[10]

[4] https://www.buergerschaffenwissen.de/
[5] https://isala.be/en/
[6] https://urbanbelonging.com/
[7] https://www.zooniverse.org/projects/victorav/the-koster-seafloor-observatory
[8] https://www.hotosm.org/
[9] https://wow.metoffice.gov.uk/
[10] https://mappingforchange.org.uk/projects/science-in-the-city-2/

### 1. *Data quality - Quality assurance frameworks*

There are many different quality assurance frameworks that local authorities and public sector organisations can draw on to assess the quality of citizen science data and whether it is fit for purpose. A good example is the UK Office of National Statistics' Quality Assurance Framework for non-official data sources.[11] The actual quality of data has significance only in the context of how it is used, meaning that for some applications, low-quality data may be acceptable. For example, in biological citizen science, many more amateur scientists can collect data over much larger areas and longer periods than would ever be possible by highly trained biologists alone. In some cases, lower quality is balanced by a wider scope, demonstrating that almost all data has value depending on the purpose for which it is used. Decision makers should define quality targets and thresholds for the minimum useful data required. This serves to not only define what data counts as accurate or to pre-define sampling approaches and protocols, but also to define when data is complete enough.

### 2. *Data governance – Data Trusts*

Issues around data governance, who uses the data and how, can be a limitation to making use of citizen science. Given the ethical imperatives around good data practices that enable open and Findable, Accessible, Interoperable, and Reusable (FAIR) data, citizen science can play a strong leadership role on data governance in the broader community of research and sustainability monitoring. One promising model for data governance is the concept of data trusts.[12] Data trusts are legal structures that provide independent stewardship of data. They are a useful way of increasing access to data while retaining trust. The organisations that collect and hold data permit an independent institution to make decisions about how that data is used and shared for an agreed purpose. The data trust becomes a steward of the data, taking responsibility to make decisions about the data and ensure they support the data trust's purpose.

### 3. *Local data expertise - Accessible resources and capacity building*

The main barrier to using citizen science data is local data expertise. There are some organisations and guidelines that provide useful resources for using non-traditional data. At a global level, the Global Partnership for Sustainable Development Data has published useful guidelines and practical resources for working with citizen generated data.[13] At an international level, the Ghana Statistical Service has paved the way in terms of building local citizen science data expertise by partnering with key stakeholders at a national and global level, including the local citizen science groups operating in Ghana and the UN Environment Programme to assess the feasibility of leveraging existing citizen science data for SDG monitoring and reporting, and for addressing the policy needs in the country.[14]

### 4. *Data infrastructure and interoperability - Data sharing platforms and data standards*

Supportive infrastructure - such as data-gathering tools, data analysis and visualisation tools, and platforms for data hosting and archiving do exist for citizen science. These include platforms such as Zooniverse, i-Naturalist, eBird, the and the Global Biodiversity Information Facility.[15] However, more work needs to be done to ensure these infrastructures are maintained and used, alongside investment in new infrastructure. Citizen science associations and other networks offer capacity building on interoperability, as well as support to facilitate the adoption of data and metadata standards. Achieving greater interoperability enables citizen science data to be more easily reused by different stakeholders, such as volunteers, researchers, and decision makers. It also means it can be combined with datasets of different scales (local, municipal, regional, national, global); and combined, exchanged and used together with different types of data. Hosting citizen science data on government portals can significantly broaden the scope and coverage of those portals, adding depth and context in sectors where government data exists, and filling gaps in sectors where it doesn't. The inclusion of citizen science data in these portals also implies that the data meets certain thresholds for methodological rigour and sustainability, either prior to inclusion, or through data cleaning and institutional arrangements implemented during the inclusion process. Doing so also can also improve public perceptions of government data initiatives.

### 5. *Lack of coherent strategies for data sharing - promoting multi-stakeholder dialogues*

There is still no coherent strategy for the development of citizen science initiatives and at a time of financial pressure, it is not possible to maintain so many different data platforms. Government agencies are working together to share information and expertise, and are also working with NGOs to develop a data sharing framework to collate and combine

[11] https://www.ons.gov.uk/economy/environmentalaccounts/methodologies/uksustainabledevelopmentgoalsuseofnonofficialsources

[12] https://theodi.org/news-and-events/blog/odi-data-trusts-report/

[13] https://www.data4sdgs.org/resources/choosing-and-engaging-citizen-generated-data-guide

[14] https://doi.org/10.1007/s11625-023-01402-4

[15] https://www.zooniverse.org/; https://www.inaturalist.org/; https://ebird.org/home; https://www.gbif.org/

data from a wide range of sources. Given the range in scales and the different topics that citizen science initiatives focus on globally, a one-size-fits-all strategy to data use will not work. For medium to large-sized data generating citizen science projects, efforts are being made to develop standards and to incorporate citizen science data into global research data infrastructure.[16] Equal efforts need to be made to enhance the utility and uptake of citizen science data into evidence for policy. As a first step, local decision makers, civil society organisations and citizen science projects should establish dialogue, map out common interests and agree on methods or standards of evidence. The Guidelines developed by the IMPETUS project on localising the Sustainable Development Goals set out a step-by-step guide for how to initiate these processes.

## Next Steps

Citizen science data can help to address the needs of decision makers as demonstrated by the benefits and opportunities listed above. However it is crucial to note that involving people in data collection and the methods for monitoring sustainability targets leads to a deeper engagement of the public in particular issues, such as marine litter and plastic pollution. This can then lead to action on the part of those involved and to a change in behaviour as a particular topic is delved into and explored in more detail. Recommendations for next steps include taking stock of data so that local decision makers can not only consider their internal use of data, but also how the public can work with it to gather new insights.

## Project Identity

| | |
|---|---|
| **Project name** | IMPETUS |
| **Author** | Alexandra Albert, IMPETUS Policy Lead, Centre for Collective Intelligence Design, Nesta, London, UK alexandra.albert@nesta.org.uk |
| **Consortium** | Ars Electronica, Linz, Austria<br>European Science Engagement Association, Vienna, Austria<br>King's College London, London, United Kingdom<br>Nesta, London, United Kingdom<br>Science for Change, Barcelona, Spain<br>T6 Ecosystems, Roma, Italy<br>Zabala Innovation Consulting, S.A., Navarra, Spain |
| **Funding scheme** | IMPETUS is funded by the European Union's Horizon Europe research and innovation programme under grant agreement number 101058677. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. |
| **Duration** | July 2022 – June 2026 (48 Months) |
| **Website** | https://impetus4cs.eu/ |
| **Further reading** | Citizen science data to track SDG progress: Low-hanging fruit for Governments and National Statistical Offices<br>Advancing sustainability together? Citizen-generated data and the Sustainable Development Goals<br>Choosing and engaging with Citizen-Generated Data: A guide<br>Citizen Generated Data and Governments: Towards a Collaborative Model<br>Mapping the landscape of data intermediaries: Publications Office of the European Union<br>Editorial: Open Citizen Science Data and Methods |

[16] This is already happening with ornithological data collected by the eBird platform which is deposited in the Global Biodiversity Information Facility which is a regular member of the World Data System.