

## 9. Citizen science: big data for environmental policies

### 9.1 Focus of the use case

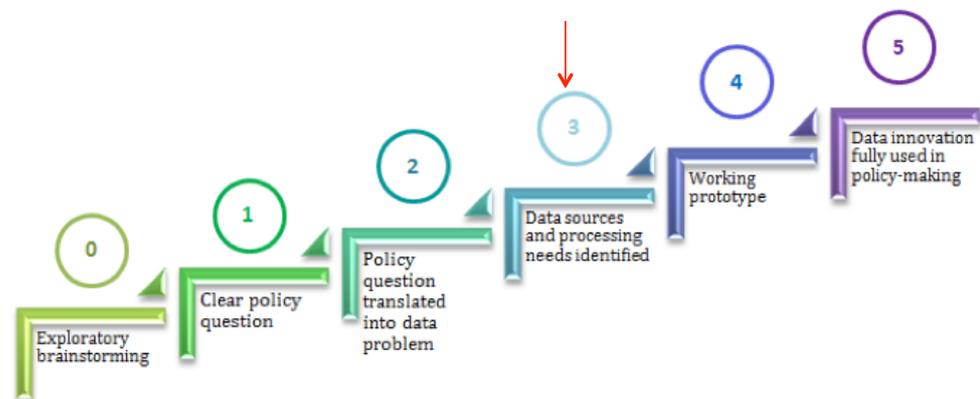
Global warming and environmental issues are at the heart of not only European Union policies, but also the UN, Organisation of American States and many other international governance bodies. The use of data is crucial in order to plan strategies not only at the global level, but also the national and regional level.

In the last decade citizen science has become an important source of data for scientists and policymakers. The definition of citizen science is scientific research that is done or curated by non-professional or amateur scientists. It is sometimes known as public participation in a research activity or crowdsourced science.<sup>66</sup> Crowdsourcing is one of the main ways to collect data in citizen science, but other open sources of data be can used in the process such as open government data and social media data.

Ideally citizen science based projects can create data-driven policies if they are widely accessible, both to the civil society and decision-makers, and not only limited to micro-scale projects. Moreover, online systems can help to create better reporting to both the public and decision-makers. This may require an element of coordination and community building between the policymakers, active users and other citizens and stakeholders.

This case study will look at several Horizon 2020 projects that use citizen science and will explain how they can help shape policies in the area of the environment. Because several projects have clarified which data sources and data analytics are relevant, we assessed the readiness level at 3.

Figure 17 Use case readiness level



### 9.2 The rationale

The EU acknowledges that data is the key in creating better environmental policies. However, it is also aware of the difficulty in gathering good quality and relevant environmental data. Data systems focus on the whole chain of environmental information, from data collection by monitoring networks - to processing and analysis of data. It also covers dissemination and communication of that data, including to the general public. We will discuss three of these major datasets in the environment field:

- **INSPIRE** - an effort to integrate data from public bodies to form an environmental map of Europe.

<sup>66</sup> Irwin, Alan. Citizen science: A study of people, expertise and sustainable development. London: Routledge, 1995.

- **SEIS** - the Shared Environmental Information System, a Europe-wide initiative to improve the sharing and re-use of environmental information (this was archived in 2014).
- **COPERNICUS** - (Global Monitoring for Environment and Security) - an initiative to combine data from land, sea and air monitoring stations with data from Earth Observation satellites.

In addition to these data collection projects, the EU identifies many international and national sources as valuable for policymaking. Yet official government data that has high quality and unified standards for EU Member States is a challenge that the European Commission is trying to address. The Commission has also identified citizen science as data that can enrich current datasets and give a better overview on the current issues and problems. The Commission has funded several policy papers on citizen science<sup>67</sup> and its connection to policy, and it funds citizen science projects as part of Horizon 2020.

### 9.3 The policy context

The European Commission is tackling the challenge of reducing emissions and sustaining healthy eco systems across Europe and the world:

*“The EU policies in the theme of environment aim to protect, preserve and improve the environment for present and future generations, proposing and implementing policies that ensure a high level of environmental protection and preserve the quality of life of EU citizens.”*

An important responsibility of the Commission is to develop and implement the specific environment action programme under the General Union Environment Action Programme to 2020 (7th EAP). The 7th EAP is referred to as the 'Living well, within the limits of our planet guidelines for European environment policy until 2020.'

The Commission works with national governments, national agencies, academia, the business sector and NGOs to increase awareness, improve monitoring and address environmental issues. The key issues of the 7th EAP include Natural Capital, Green Economy, Environment and Health, EU regulation, and global challenge of knowledge related to these issues. As part of its strategy, the EU promotes legislation and directives on energy, climate change and protecting wildlife. In addition, the Commission can prosecute and bring to justice Member States that do not follow EU directives. Finally, the EU funds various research programmes under the Horizon 2020 programme for research and innovation on the environment.

Citizen Science is promoted as part of the open science dossier of the Digital Agenda for Europe.<sup>68</sup> In terms of policymaking it has three main goals:

- Efficient and transparent use of public and private science and research funding,
- Better engagement with research, governance and accountability,
- Bringing European policymaking closer to the people while basing it on scientific evidence.

In terms of the policy cycle, citizen science can be found in various stages of the process. It may help to raise awareness through participation of citizens and their

<sup>67</sup><https://www.wilsoncenter.org/publication/citizen-science-and-policy-european-perspective> ; <https://ec.europa.eu/digital-agenda/en/news/green-paper-citizen-science-europe-towards-society-empowered-citizens-and-enhanced-research-o>

<sup>68</sup> <https://ec.europa.eu/digital-agenda/en/citizen-science>

engagement with the topic. Another path is to use the data that was collected by citizens to define a problem and/or to monitor how environmental issues are evolving.

The Commission has produced a green paper about citizen science and policy, in collaboration with Socientize, the main EU portal for citizen science.<sup>69</sup> It also produced a specific paper about citizen science for environmental proposes.<sup>70</sup>

In the field of environment, the EC funded several research projects based on citizen science. The set of projects includes:

- **Everyaware** - A project that uses citizen science to collect data on air quality and noise.<sup>71</sup>
- **GAP2** - A project that demonstrates the role and value of stakeholder driven science within the context of fisheries' governance.<sup>72</sup>
- **BIG4** - A global consortium to train a new generation of systematic entomologists with broad profiles amalgamating classical aspects of biosystematics of the four biggest order of insects, including cross-disciplinary methodological innovations.<sup>73</sup>

## 9.4 The data process: from data collection to analysis and visualisation

### Data Sources

As mentioned before, citizen science often uses volunteers to collect data (cf. crowdsourcing). This requires not only finding volunteers, but sometimes also making sure they cover different geographic areas in order to get a good sample. In addition, volunteers can be divided into different groups – targeted, opportunistic, stationary or mobile. Targeted volunteers means that volunteers need to report on specific events; they need to be guided and to pay attention to what is done around them. The opportunistic group does not need to actively report what was done around them, the sensor or any other tool will do it for them. This group can use for example, social sensing tools or a dedicated GPS sensor on their phone that tracks their activities. In addition, some volunteers have stationary sensors while others carry the sensors with them.

Volunteers can come from different groups - university students, young moms, elderly, etc. - and may need to be instructed thoroughly. Volunteers come with different knowledge about the technology or the environment and have to be guided before performing measurements. In some cases, there is no control over the sample, and the geography is based on where the volunteers are rather than on a representative distribution.

To get a sufficient sample there is often a need to obtain targeted populations. It has been suggested that neighbourhoods or schools could be targeted, but it might also be useful to target environmental or other organisations or to focus on an online community that is more technologically oriented to perform measurements. The main challenge in citizen and crowdsourced data, especially when it is focused on events, is to keep volunteers doing measurements for long periods of time. These volunteers need to be rewarded for their time or be incentivised to carry out this work. These incentives may not be sustainable and so need to be given due consideration. School children can be very helpful since they can carry out these measurements as part of their school assignments. It may also be helpful to focus efforts on areas that suffer

---

<sup>69</sup> Green paper for citizen science - <https://ec.europa.eu/digital-agenda/en/news/green-paper-citizen-science-europe-towards-society-empowered-citizens-and-enhanced-research-o>

<sup>70</sup> Science for Environment Policy, Environmental Citizen Science  
[http://ec.europa.eu/environment/integration/research/newsalert/pdf/IR9\\_en.pdf](http://ec.europa.eu/environment/integration/research/newsalert/pdf/IR9_en.pdf)

<sup>71</sup> <http://www.everyaware.eu/the-everyaware-project/>

<sup>72</sup> <http://gap2.eu/>

<sup>73</sup> <http://big4-project.eu/>

from poor air quality or noise quality where there is a demand for changing these conditions.

The data should ideally be owned by the people who collected it, as well as open and free to use with an easy interface for analysis. If so, policymakers can make use of the data as part of a transparent policymaking process and also aim to enrich existing data sources (and to make explicit when traditional data sources conflict with data obtained by citizens).

There are different ways to collect the data. Some can use dedicated smartphones apps to collect data either as a passive sensor or as an active action whereby the volunteer enters data. There are also open source solutions that are platforms for data collection such as Pybossa, which allow any researcher to create a free and secure data collection portal where citizens can enter their findings.<sup>74</sup> Another interactive way to collect data is to add a gamification element to it.<sup>75</sup> This can be done by adding elements such as scoring or badges to prompt participation from users who feel they are more in a game than a research project. Galaxy Zoo uses some of these elements on its website.<sup>76</sup> The Everyaware team had an innovative idea of using the data through a game platform that will allow users to learn about the environment and their surroundings using the data they collected. It seems that there was a working prototype, but so far the use of the data was quite basic and was done in one language only.

### Data analytics and visualisation

Citizen science projects use a variety of ways to visualise their data. In the data collection processes, visualisations are often used in mobile apps, so the users can see their tracks on a map, calculate the estimated black carbon exposure, and follow sensor output in real time plots. In policy processes, projects use different approaches such as heat maps and standard graphs such as pie charts to visualise the data.

Another way in which data is analysed is through open datasets. These datasets enrich citizens' research and allow anyone to analyse the data. If combined or cross referenced with crowdsourced data, this can be a valuable tool for policymakers.

## 9.5 Reflections on challenges and next steps

Environmental issues are among the top priorities of the EU and other international and national policymakers. Financial resources are invested in research projects and collection and organisation of data that will allow development of new tools for a cleaner environment. This will not only allow policymakers to sustain and preserve nature and its inhabitants, but also improve the quality of life in the EU and among a global population.

Data on environment is collected by governments, agencies and academic institutions from around the EU and the world. These data, after analysis that is usually done by experts, can assist policymakers in formulating strategy in particularly complicated area of policy. Citizen science has been a major approach in research in this area, and it is innovative in the policy cycle since it uses citizens to collect data, analyse it, and, in the future, may be used to formulate policy that is closely tied to citizens' geographical locations.

Privacy requires attention. Some projects use sensors which track movements and routines of participants. If not handled correctly, this data can risk violating participants' privacy. In addition, the sensors use the participants' mobile phones as hardware. This kind of software can, in theory, access participants' private data on

---

<sup>74</sup> <http://pybossa.com/>

<sup>75</sup> Bowser, Anne, D. Hansen, and J. Preece. "Gamifying citizen science: Lessons and future directions." Workshop on Designing Gamification: Creating Gameful and Playful Experiences. 2013.

<sup>76</sup> <http://www.galaxyzoo.org/>

their mobile. There is a need to be vigilant with the data and make sure these data breaches will not happen.

Moreover, it is important to remember that the majority of citizens will not necessarily have the resources or knowledge to analyse such data, nor have the means to use them to improve their quality of life. As such, there is a risk that data availability will differ between rich and poor areas, between urban and rural areas, etc.

While citizen science applications are on the web, they are not easy to find, and users need to have the technical knowledge to engage with the applications. In addition, mobile applications will be likely used by people who are tech savvy or with knowledge of how to use a smart phone, and exclude other populations such as the elderly, who are less likely to use such applications.

Using citizen science means sustained engagement with citizens, including large parts of the population that have limited scientific expertise. In order to create real engagement with the general public, there is a need for an educational element. The population should understand the benefits of the project in order to participate.

In addition, research shows that even in the internet age, civic engagement is still low (Zuckerman, 2014). In order for this engagement to be effective, there is a need for long cycles of measurement that require commitment from users. To scale up projects, there is also a need for solutions to keep users engaged over time. It has been suggested that the education system is a place to implement longer cycles of engagement, but this will also require sustained cooperation from the public officials and resources for training.

There is an assumption in research that engaged users will be able to advocate for better quality of life on environmental issues. For effective advocacy with policymakers, however, more is needed than just measurement and data visualisation. Better approaches for the use of data in the policy cycle are a crucial precondition for citizen data to be influential in the policy cycle.

### **Further reading**

Commons Lab. Citizen Science and Policy: A European Perspective. Wilson centre, 2015. <https://www.wilsoncenter.org/publication/citizen-science-and-policy-european-perspective>

Cohn, Jeffrey P. "Citizen science: Can volunteers do real research?" *BioScience* 58.3 (2008): 192-197.

Irwin, Alan. "Citizen science." London: Routledge, 1995.

Silvertown, Jonathan. "A new dawn for citizen science." *Trends in ecology & evolution* 24.9 (2009): 467-471.

Zuckerman, Ethan. "New Media, New Civics?" *Policy & Internet* 6.2 (2014): 151–68.