Citizen-generated data within the Green Deal data space

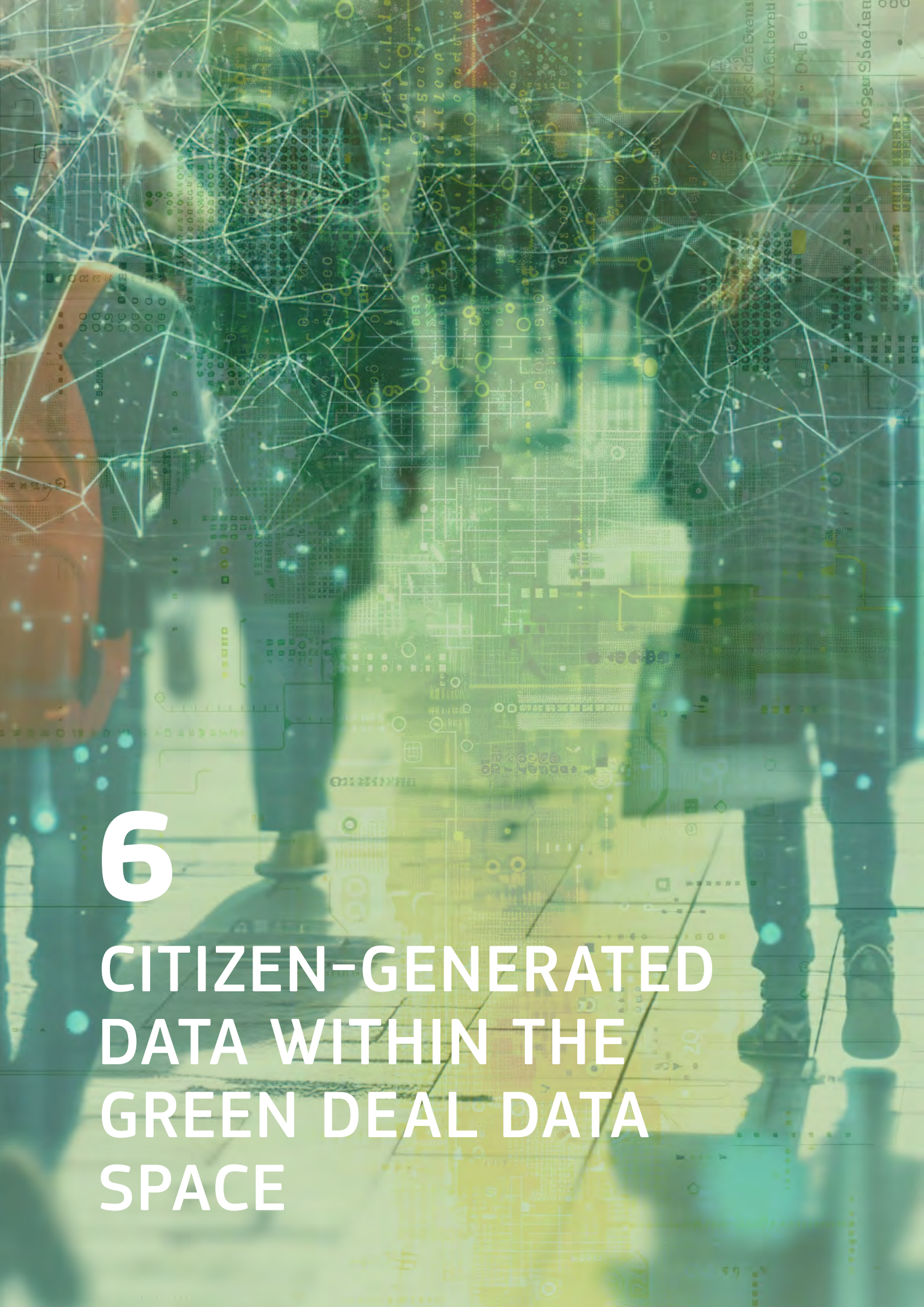
Milan, S.

DOI
10.2760/2422339
10.2760/0517622

Publication date
2024

Document Version
Final published version

Published in
Unlocking Green Deal Data



6

CITIZEN-GENERATED DATA WITHIN THE GREEN DEAL DATA SPACE

We are at a turning point in history. On the one hand, disruptive digital innovation and the rapid growth in computing power have resulted in unprecedented data capabilities. On the other hand, environmental depletion is at its highest, and the consequences – from the climate crisis to land degradation – touch all dimensions of people's lives. This chapter explores how the positive consequences of the 'data revolution' can be harnessed for the ecological transition. It examines how individuals and communities, including laypersons, can generate and (re)use data on environmental issues of concern to drive changes in social norms, laws, and regulations and inform policymaking.

In today's increasingly data-driven society, a vast array of tracking and sensing technologies is transforming human and social activities into data points. Never in history has humanity possessed such extensive knowledge about itself and its environment (Cukier and Mayer-Schoenberger, 2013). Data production and analysis are enabled by software and computing infrastructure, including cloud capabilities and data centres. Well-resourced industry players take the lion's share in the data race, operating at a large scale for profit-driven purposes such as the monetisation of personal data for marketing and individual profiling (Couldry and Mejias, 2019). Their activities have a staggering environmental footprint (Monserrate, 2022).

Countering these trends, a distinctive type of data has emerged – citizen-generated data (CGD). Thriving on the fringes of the mainstream data ecosystem, CGD benefit from the recent availability of accessible, low-cost tools, and is often the result of participatory collaborative initiatives. Rather than 'big data', CGD are frequently 'just good enough data' (Gabrys, Pritchard, and Barratt, 2016, p.2), and challenge 'standard practices used by regulators' (Ottinger, 2010, p.245). They provide valuable opportunities to enhance environmental governance and foster citizen participation and empowerment in the pursuit of sustainability. Consequently, CGD can play a significant role within the framework of the EU Green Deal data space (GDDS), in connection with the opportunities and tools provided by the new legal instruments underpinning the GDDS. For example, the definition in the Data Governance Act (DGA) of trusted data intermediaries (European Union, 2022) paves the way to empower citizens' use and management of their own data, and the creation of personal data spaces (Lähteenoja, 2023). Furthermore, the Data Act (European Union, 2023) provides additional data access rights to users of connected devices and services, such as smart watches, home supplies, etc. Lastly, the Open Data Directive (European Union, 2019) makes public sector data freely available for reuse, opening the opportunity for citizens to access and use these data.

This chapter delves into the nature, challenges and opportunities of CGD in the context of GDDS. It has five objectives:

1. Providing an actionable definition of CGD within GDDS.
2. Positioning citizens as knowledge makers, and knowledge held, generated, or supplied by citizens as complementary to academic and professional expertise, and on equal footing with it.
3. Identifying the main challenges encountered by individuals and groups engaging in data generation and data use in the context of CGD in environmental governance.

4. Reviewing policy measures to enable, foster, and support the generation and use of CGD.
5. Providing a set of policy recommendations ('necessary measures') to create an enabling environment for CGD in the context of GDDS.

The chapter is organised as follows. Firstly, it presents the methodology underpinning the research. Secondly, it reviews the conditions that enabled the emergence and growing popularity of CGD. Thirdly, the Findings section opens with an actionable definition of CGD and recognises citizens as legitimate knowers, essential for maximising the potential of CGD in environmental governance. It then offers an action-based concept map to enable policymakers and CGD actors to classify the core activities and phases of the CGD lifecycle. The chapter reflects on challenges and open questions and explores the features of an enabling ecosystem for CGD. Finally, it proposes a set of necessary policy measures to support CGD supply and utilisation, leveraging its potential for the benefit of society as a whole.

6.1 Methodology and data

The chapter answers three main questions:

- What are CGD comprised of and how can such data contribute to promoting environmental awareness and a just, sustainable transition?
- What challenges are encountered by CGD?
- What does an enabling environment for CGD look like? What measures can be kick-started or scaled up, and how can they be encouraged?

Data for this chapter are derived from a qualitative analysis, capable of generating 'thick' evidence (Geertz, 1973) on the problem at hand:

- **Desk research:** This includes academic literature across disciplines, as well as policy papers and grey literature by stakeholders like industry and civil society organisations.
- **Interview and ethnographic data:** Collected within the framework of the DATACTIVE project (2015-2021, PI S. Milan, European Research Council StG no. 639379, data-activism.net).

The term 'citizens' in this context refers to individuals and communities who are not professional data producers or analysts (e.g. laypersons) and/or do not have a codified role in data handling (e.g. those outside the realm of the State and industry). In the context of CGD, the reference to citizenship is detached from its legal definition as the bond between an individual and their State. Here, 'citizens' is a synonymous for 'people', also including those who do not hold citizenship rights.

Finally, this chapter adopts a global perspective to identify commonalities in CGD across diverse contexts. Despite minor socio-cultural differences, the nature and dynamics of CGD are primarily determined by the combined effects of data and data infrastructure on the one hand, and an enabling environment, on the other. These are the primary concerns of this report.

6.2 Background

As society increasingly turns into data ('datafies') various aspects of human activity – from personal health and interpersonal connections to public safety and social security – citizens and institutions alike are becoming more aware of the crucial role of information in democratic systems (Hintz, Dencik, and Wahl-Jorgensen, 2018). This 'datafied society' represents a fundamental paradigm shift in

how we understand and manage the policy and the public good (Kitchin, 2014). However, it is characterised by an unparalleled power asymmetry between the State and its citizens (Brunton and Nissenbaum, 2015). The advance of generative artificial intelligence (AI) augments this power asymmetry.

While today's data-driven environment is often linked to disempowerment rather than democratic agency (cf., Couldry and Powell, 2014), transformative experiments have emerged where citizens repurpose data in innovative ways. *InfoAmazonia* (infoamazonia.org) is one of many examples showcasing the unprecedented potential for laypersons to drive social change through active engagement with data generation and analysis. Since 2012, this network of citizens and journalists from the eight countries of the endangered Amazon region has been leveraging crowdsourced, remote sensing and satellite data to produce investigative reports, maps, and advocacy material. Promoting data transparency in the Amazon through CGD helps to prevent deforestation and wildfires (Milan and Gutiérrez, 2017).

Data from such initiatives are referred to as citizen-generated data. CGD can enable individuals to exercise political agency in a data-driven society (see Milan, 2018). It includes both data that are passively contributed (e.g. cellular phone tracking) and data that are actively generated (Haklay, 2013). When actively generated, CGD often result from a form of political engagement that uses data and data generation techniques as opportunities to drive policy change or societal transformation, known as 'proactive data activism' (Milan and van der Velden, 2016).

The origins of CGD trace back to citizen science initiatives and align with the role of technological innovation in democratising data collection processes. Early manifestations of CGD were often manual and localised, but the penetration of user-friendly digital tools in society, such as smartphones and low-cost sensors, has expanded the scope and scale of data collection efforts. Accessible data analysis and data visualisation software have further facilitated this expansion (Gutierrez, 2018). Interestingly, the affordances of a given technology – the possibilities and limitations it offers to its users (Davis and Chouinard, 2016) – shape how individuals can interact with such technology and what actions are possible, also in relation to the users' abilities and goals, and mould the societal dynamics it triggers (cf., Baack, 2018).

The growing interest in CGD goes hand in hand with the evolving societal perception of what constitutes legitimate knowledge at a given point in time and how it is achieved. Gray, quoted in Kitchin (2014), argues that science has entered a fourth paradigm, predicated on the increased availability of new data and new analytics. This fourth paradigm is a type of 'exploratory science' that is data-intensive and rooted in statistical exploration and data mining. Haklay et al. present citizen science, which involves non-experts in data generation, as a valid tool in 'post-normal science' – a framework 'concerned with the social robustness of applied science, science-based professional consultancy, and scientific advice for policy in situations of high stakes, high uncertainties, and contested values' (Haklay et al., 2023, p.1). Amidst these developments, **CGD can be seen the manifestation of a new way of engaging in knowledge generation**. Its disruptive potential has reconfigured how evidence is produced within the realm of participatory governance.

In the context of environmental and climate governance, CGD are frequently mobilised

as part of environmental justice initiatives, especially when they include contestation over knowledge claims from the ground-up (Ottinger, 2024). CGD can contribute to 'environmental democracy, a set of principles and practices that allow people to participate in environmental decision-making processes that affect their lives' (Berti Suman et al., 2023). Environmental democracy was first enshrined in the Convention on *Access to Information, Public Participation in Decision-Making, and Access to Justice in Environmental Matters* (also known as the Aarhus Convention, 1998) adopted by the United Nations Economic Commission for Europe (UNECE). It is grounded on three key pillars deemed essential for empowering people and communities to protect their environment and to hold governments and other entities accountable for their actions affecting the environment:

- Openness and transparency in environmental governance, including access to information.
- Participation by the public in environmental matters of concern.
- Access to justice in environmental matters and decision-making.

The Commission has also addressed the importance of CGD for the environment, as highlighted in the 2020 Staff Working Document on 'Best Practices in Citizen Science for Environmental Monitoring', which provides key insights on the opportunities for and benefits of using citizen science for environmental monitoring (European Commission, 2020b). Examples of CGD deployed within environmental justice initiatives include a bottom-up initiative monitoring the noise levels of Amsterdam Schiphol airport (Netherlands), where citizens autonomously measured noise levels of manoeuvring airplanes with microphones to challenge the 'informational monopoly' of the State and reclaim their right to live in a healthy environment (Berti Suman, 2018). Similarly, in the US, in oil fields subjected to 'fracking' (an invasive drilling technique used for extracting oil or natural gas from deep underground), professional researchers have developed monitoring kits for residents to track air pollution, one of the byproducts of fracking, thus evidencing harm (Gabrys, 2017). In the riverbed of the Rio Magdalena, one of Colombia's largest watercourses, a group of fishermen concerned about the impact of large-scale gold mining created a time-series of riverbed transformation, through drawing on paper and large canvases. The effort was aimed at advocacy in the absence of updated satellite imagery (author's fieldwork, 2017). But people may also mobilise to defend government-generated data at risk of deletion. Following the 2016 election of climate-denier Donald Trump to the US presidency, citizens organised 'DataRescues', harvesting, scraping and archiving federal websites and datasets related to climate change in fear that they would be removed (Currie and Paris, 2018). It was a way to engage in 'data resistance' (Vera et al., 2018).

These cases demonstrate the growing public interest in CGD to support environmental knowledge and policymaking, acknowledging its potential to 'contribute to the framing of or offer a different perspective on an environmental issue they care about' (Berti Suman et al., 2023, p.2). Despite the democratic potential of these citizen-led responses to the environmental crisis, CGD faces obstacles when trying to give teeth to the knowledge it generates. One such obstacle concerns the recognition of the legitimacy of citizen-generated knowledge. This is because citizen science approaches 'often do not fit within current institutional practices' prevalent in policy cycles (Schade et al., 2021, p.361). Scientists'

preferences might also hinder the adoption and limit the applications of citizen science data, due to the scarce awareness of citizen science projects and dynamics, a persisting distrust in citizen science data quality, and a bias favouring data collected by fellow scientists (Burgess et al., 2017). Moreover, people from marginalised communities or identities typically suffer from systematic 'exclusion from judgment, inadequate epistemic resources, and denial of status as knowers' (Ottinger, 2024, p.199).

Effective promotion and deployment of CGD depends on society's ability to recognise it as a valid form of knowledge generation. A renewed lexicon is needed, which is presented next.

6.3 Findings

6.3.1 Key terms and definitions in citizen-generated data

This section proposes a renewed lexicon for CGD, assuming that the way we imagine and discuss the nature and role of CGD influences their likelihood to impact societal transformation and policymaking. First, it provides an actionable notion of CGD. Second, it posits individuals and communities as knowledge makers capable to accelerating the ecological transition.

BOX 3.

Actionable definition of CGD.

CGD indicate the diverse range of information that:

- is created or collected, analysed, reused, and/or disseminated by non-specialists, including both individuals and community groups or thematic collectives (i.e. mobilising on a specific matter of concern);
- independently or in collaboration with specialists, including scientists or subject-matter experts as well as governmental and non-governmental organisations;
- often but not necessarily leveraging digital technologies;
- either not-for-profit, or in combination with for-profit activities; and
- adheres to ethical guidelines and privacy protections to ensure responsible and respectful data use.

CGD are generally ascribed to the realm of 'non-traditional data sources' (Fritz et al., 2019). Two sets of descriptors provide a structured way to define and understand the various aspects and applications of CGD. These descriptors are not mutually exclusive.

In relation to the nature of CGD, we can identify five sets of descriptors, whose combination allows for labeling information generated by citizens in the context of GDDS:

- **Data types:** Quantitative data (numerical data such as counts, measurements, statistics), qualitative data (descriptive information in text, photos, videos, audio recordings), spatial data (georeferenced data, localised data), temporal data (capturing changes over time).
- **Data collection methods:** Sensor-based data (generated using sensors to provide measurements), crowdsourced data (information collected by many

contributors via platforms or apps), remote sensing data (generated using drones or other remote sensing technology), manual observations (through human scrutiny such as reporting, journaling, etc.).

- **Tempo:** Real-time data, historical data, time-series, etc.
- **Engagement level:** Individual, community or collective, collaborative (leveraging multiple sources of observation and expertise); passively generated data.
- **Data quality:** Raw data (unprocessed information), processed data (cleaned, validated, and analysed), or verified data (cross-checked for reliability).

Citizen-generated data can contribute to the understanding, monitoring, and management of environmental issues. Simultaneously, it typically aims to promote transparency, inform policy, and/or foster community engagement. In relation to its purpose and use, we can identify four main orientations in CGD:

- **Policy-oriented:** When data availability and policy formulation are enhanced by CGD. For instance, CGD can supplement official data and support policy development and implementation.
- **Process-oriented:** When transparency is enhanced by CGD. For example, CGD can promote transparency and accountability in environmental decision-making.
- **Outcome-oriented:** When adaptive management is enhanced by CGD. For example, CGD can improve the responsiveness and adaptability of institutions and foster community resilience.
- **Public-oriented:** When public engagement is enhanced by CGD, encouraging awareness raising, promoting literacy, and boosting public participation, thus empowering citizens to actively engage in environmental governance.

6.3.1.1 Citizens as knowledge makers

Individuals and groups are positioned by CGD as knowledge makers, capable of highlighting on-the-ground realities that might otherwise be overlooked by institutional efforts. They can identify and expose the community's experience of a given problem, including quantified measurements as well as actionable evidence on perceptions, fears, and needs. In doing so, they leverage and become a proxy for a type of expertise derived from lived experiences rather than academic or professional qualifications: **experiential knowledge**. Experiential knowledge can complement academic and professional expertise, offering a holistic view of complex issues and contributing to more inclusive and effective policymaking.

The recognition of citizens as knowledge makers is supported by a growing body of literature, particularly within science and technology studies (STS). Concepts such as 'boundary work' (Gieryn, 1983), 'co-production' (Jasanoff, 2004), 'street science' (Corburn, 2005), and 'civic technoscience' (Wylie et al., 2014) have emerged to problematise the sources of scientific expertise and underscore the importance of experiential knowledge as a legitimate and powerful form of expertise. However, experiential knowledge often lacks recognition in policy cycles or society (Ottinger, 2010). The problem originates from how the category of 'expert' is defined and the public trust attributed to it. As a result, 'alternative, more culturally rooted and legitimate forms of collective, public knowledge – and of

corresponding public order – which could arise from the informal non-expert public domain are inadvertently but still systematically suppressed' (Wynne, 1998, p.46). This may be ascribed to the tension between expertise and experience, which can be addressed by distinguishing between distinct types of expertise, including 'interactional expertise', whereby laypersons gain sufficient understanding of an issue to meaningfully engage with scientific discourse (Collins and Evans, 2002). Importantly, knowledge produced by citizens is 'not only local, situated and experiential in origin but also *collectively generated and held*' (Scott, 2016, p.261 emphasis added).

The sociology of social movements has further emphasised this collective dimension of knowledge-making. Grassroots movement actors can create 'collective spaces of knowledge production' that 'foster the coordination of disconnected, local, and highly personal experiences and rationalities within a shared cognitive system', supporting collective claim-making (Della Porta and Pavan, 2017, p.297). Movements are the harbingers of subaltern knowledge, opposed to official knowledge, where the standpoint from which knowledge is elaborated matters (Cox and Flesher Fominaya, 2009). They focus on power inequalities, the realities of vulnerable communities, and the role of reflexivity in the making of knowledge (Ryan et al., 2010).

A data justice perspective, explicitly integrating social justice concerns into the analysis of data-driven societies (Dencik and Sanchez-Monedero, 2022), stresses the meaningful participation in knowledge generation of all stakeholders, especially those impacted by data practices. Taylor (2017) identifies three pillars of data justice: visibility, engagement with technology, and non-discrimination. These pillars highlight crucial elements such as access to representation, autonomy in technology choices, and the capacity to challenge data biases, which speak to the role of citizens as knowledge producers. In environmental governance, this approach helps counter the 'extractive logic' of the State when it separates data from its provenance (Vera et al., 2019).

The question that remains largely unaddressed in these accounts is that of legitimacy, which can be defined as the condition whereby an actor is regarded by others as exercising its knowledge-making power in an authoritative and appropriate manner (Berti Suman et al., 2020). Legitimacy is not just contingent backing, where citizens as knowers are granted limited, ephemeral support in certain policymaking processes. At a more fundamental level, it involves trust (in the quality of the knowledge produced), and confidence (in the dynamics of knowledge production) (adapted from Scholte, 2019). Acknowledging the role of individuals and communities, including laypersons, as knowers means reclaiming and reinstating the legitimacy derived from experiential rather than merely credentialed knowledge. The next section explores the lifecycle of CGD, which contributes to elucidate why citizens are legitimate knowers.

6.3.2 The lifecycle of CGD: An action-based concept map

Analysing the ideal-type lifecycle of CGD is crucial to understanding the dynamics of CGD generation and use. The lifecycle of CGD includes six stages, illustrated in Figure 1, which presents an 'action-based concept map' representing the relationships between the stage of CGD, with each stage being described by an action verb. This is in line with the notion of **data as enabler** and CGD as enabling a set of activities (namely, activism, advocacy,

solution development), but also transformations (citizen empowerment, behavioural change, and citizen agency). The map is illustrated with examples from the realm of environmental and climate governance.

There are two modes of **making data**, depicted on the left-hand side of the map. Citizens can

- **Collect existing data**. Generated by utility meters, smartphones, or meteorological stations in an automated fashion, data are assembled, analysed and deployed by citizen knowers for awareness raising, evidence generation, or advocacy. For example, the EU-funded DAIAD project (2014-2017) empowered urban dwellers to monitor their data consumption by collecting readily available records to be compared with consumption figures of reference groups, with the goal of lowering water use (Sartorius, 2017). The DesPat app turns Android smartphones into a privacy-respecting camera-based pedestrian tracking tool to analyse pedestrian traffic patterns in an automated fashion (Getschmann and Echtler, 2021). But citizens can also
- **Produce new data** about matters of concern for which evidence is not yet available. The EU-funded D-NOSES project (2028-2021, dnoses.eu) is a case in point: to tackle the problem of odours, it empowered citizens to map and report odour incidents in real-time, notifying type and intensity in the *OdourCollect* app. It gathered over 10,000 odour observations from 1,600 registered users. *FreshWaterWatch* (freshwaterwatch.org) involves individuals and communities in the collection of quality data. 'WaterBlitz' events, organised at regular intervals, enable swaths of citizens to simultaneously test their local waterbody with freely distributed water testing kits.

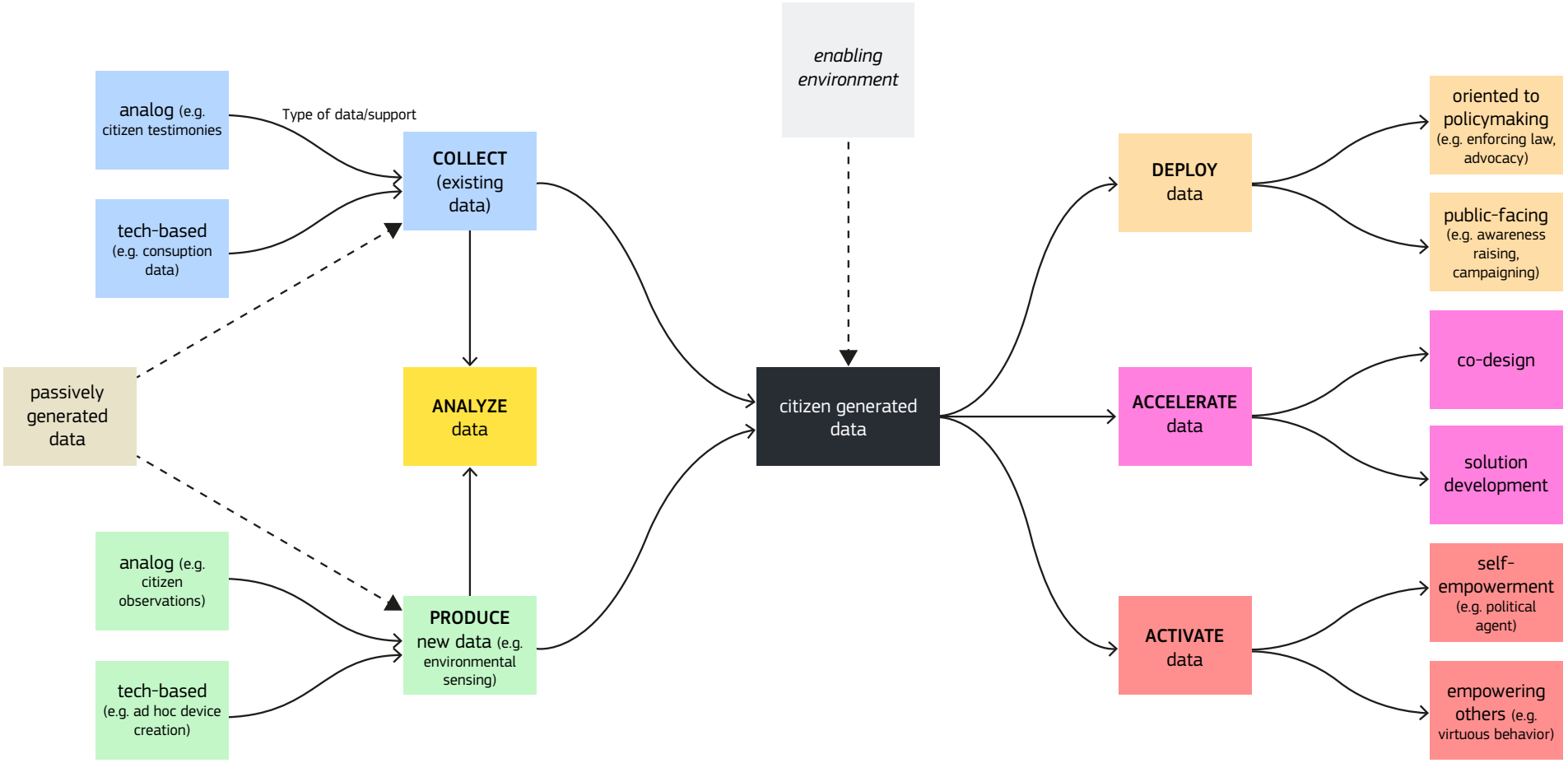
In both instances, we can differentiate between analogue data (such as citizen testimonies or logs and diaries commonly used in conservation projects) and tech-based data, including various digital and non-digital modalities. Examples of tech-based data include ad-hoc devices for environmental sensing or curated data from buckets used to monitor air toxics (Ottinger, 2010). Both analogue and tech-based data can also incorporate passively generated data.

On the right-hand side of Figure 1, we find three main modes of **exploiting CGD**. Citizens can

- **Deploy data:** CGD are curated and arranged for consumption by third parties (e.g. State, industry, public), either mobilised for informing policymaking, enforcing laws or institutional change, or for awareness raising, campaigning, and public education. This is the most popular way of leveraging CGD. For example, people monitored air pollution near fracking sites to make visible the 'experience of citizens living on the gas fields' and seek redress (Gabrys, 2017, p.177).
- **Accelerate data:** CGD are used for co-design (e.g. when citizens model policies jointly with public officials) or solution development (including product development). In Spain, the D-NOSES project accelerated citizen data to develop a national odour pollution standard (2017).
- **Activate data:** CGD are used to enhance individual practices or empower others by promoting virtuous behaviour. This approach is particularly significant in achieving the Green Deal objectives, as it encourages citizens to improve their daily actions. The EU-funded Waste4Think project (2016-2020) use data to illustrate campaigns aimed at reducing landfill waste and increasing recycling in various European towns (Waste4Think, 2020).

FIGURE 4.

Citizen-generated data: An action map.



Source: own elaboration.

6.3.3 CGD: Challenges and open questions

Six pressing issues represent the main weak spots in the CGD lifecycle, and might hamper the validity, legitimacy, and utilisation of CGD in environmental and climate governance. They combine public-facing and internal aspects of the CGD lifecycle.

- **Perceived data quality**, as defined by scientific norms, is often seen as the main obstacle to the legitimacy and validity of CGD (Fritz et al., 2019). The term holds different meanings for various stakeholders (Balázs et al., 2021), while policymakers require 'regulatory-quality data' (Ottinger, 2024, p.207). Moreover, data quality requirements often clash with the open, collective, and anonymous nature of CGD (Lukyanenko, Wiggins, and Rosser, 2020). Lack of standardised protocols of data collection, storage, and analysis as well as data verification and validation can hinder utilisation and the effectiveness of CGD (Ottinger, 2010).
- **Data management** means making CGD findable, accessible, interoperable, and reusable (Fritz et al., 2019). This involves challenges related to the technical infrastructure (e.g. maintenance and upkeep, limited access to storage for archiving, scalability) as well as to data integration, interoperability (ensuring that CGD can be easily integrated with official data sources and other types of environmental data), compliance with legal and regulatory requirements, and data privacy and security. These challenges require a holistic perspective that combines both technical and legal expertise.
- **Data ownership** in CGD tends to be diffuse, due to the collective nature and the frequent anonymity in data collection. The awareness often remains within the community of knowers. Indigenous knowledge is a case in point (Walker et al., 2021); the specificities of indigenous knowledge production may also question the fitness of standardised ways of eliciting consent for data handling (author's fieldwork, 2020). This diffuse ownership contrasts with the credentialed ways of producing knowledge, such as in academia or the consultancy industry. Data sharing licenses might pose a challenge, unless they account for collective ownership.
- **Sustaining engagement and participation** points to three orders of problems. CGD are typically under-resourced and based on volunteer work. Incentives are needed to sustain participation and motivation over time, making sure people contribute data regularly. Ensuring diverse and inclusive participation across demographics (inclusivity) is also a challenge, as the engagement of CGD is tied to education and access to resources. Finally, capacity building is a key ingredient: expertise, which is often individually owned, can be shared within the community.
- **Vulnerability of CGD makers**. Marginalised knowers, vulnerable individuals and communities, or people in emergency situations, including aid workers, tend to be at the forefront of CGD production (Gutierrez, 2018), but might not find themselves in the ideal conditions to supply consistent flows of high-quality data. Because 'marginalised people are often marginalised as knowers

as well', recognition and participation of citizens in data making are not sufficient conditions to support CGD (Ottinger, 2024, p.200). A data justice perspective can help bring power imbalances and the rights of vulnerable groups to the forefront.

- **Governance and policy integration**. Despite growing in popularity, CGD often lack recognition in policy and scientific circles, making effective use in policymaking uncertain (see, among others, Schade et al., 2021; Berti Suman et al., 2023). Fostering collaboration between citizens, nongovernmental organisations, government agencies, and other stakeholders requires further experimentation. Establishing regulatory frameworks to support and govern CGD in environmental governance is a key challenge, as explored next.

6.3.4 CGD: Towards an enabling ecosystem. Necessary measures to support CGD

Amidst the significant shifts in power dynamics within the 'science-society-policy interface' (Schade et al., 2021), it is crucial to explore how to empower stakeholders in CGD and enhance their effectiveness for environmental and climate governance. This section draws from the frameworks of environmental democracy, data justice, and epistemic (in)justice (Fricker, 2007) to outline the key components of an enabling ecosystem for CGD. It centres 'careful knowing' – an approach inspired by the feminist notion of care, responsive to the needs of marginalised knowers, and complementary to participation and recognition in environmental justice. At the very least, careful knowing means 'developing epistemic resources and standards of evidence suitable to the specific circumstances of frontline communities, as well as the need to care for community members' status and self-regard as knowers' (Ottinger, 2024, p.215).

Starting from a system thinking perspective, it is essential to recognise that data collection and data use are interdependent and cannot be separated. Thus, **an enabling ecosystem mediates between the creation of CGD ('making') and its application ('exploiting') in policymaking and awareness raising**. Here, the concept of an 'enabling ecosystem' goes beyond legislative frameworks (as further explained in Chapters 5 and 8 of this report) and encompasses issues of legitimacy, accountability, and empowerment. In line with system thinking, it attributes responsibility not only to institutional actors (e.g. policymakers, funding bodies), but also to the CGD community itself. By addressing these aspects, we can create an environment that supports effective CGD use in environmental and climate governance.

Six elements have been identified as part of the enabling ecosystem: at the community level, literacy and accountability; at the material level, resources and infrastructure, and at the institutional level, legitimacy and recognition on equal footing.

The first class of elements of an enabling ecosystem for CGD intervenes at the community level.

- **Data literacy as flywheel of participation**. Within CGD initiatives, experiential knowledge becomes meaningful when combined with data skills (e.g., Sander, 2020). The challenge is to 'go from something that's more of an expert culture [data handling] to transferring this to people' (DATACTIVE interview, 2019). Whilst fostering data literacy programmes would also be the task of governments and funding bodies, it rests also on the CGD community to

promote grassroots training, knowledge sharing, and mutual learning.

- **Accountability** within an enabling ecosystem for CGD refers to the mechanisms, practices, and policies that ensure all participants – citizens, researchers, policymakers – are responsible for their actions and contributions to the data collection and usage process. This would involve (a) transparency in data collection and usage, and open communication with CGD participants and data subjects; (b) ethical conduct, ensuring data are collected and used responsibly without causing harm; (c) data stewardship by those managing the data, who must ensure accuracy, privacy, and security; and (d) equitable opportunities for all, particularly marginalised knowers, to participate and influence the data processes. Commonly agreed standards and guidelines would contribute to mainstream accountability within CGD communities and initiatives. Registered Data Altruism Organisations (RDAOs), defined by the Data Governance Act as non-profit trusted entities that make relevant data available within a system of system of safeguards that protect the rights and interests of citizens and companies, have a role to play in driving this process.

At the material level, an enabling ecosystem for CGD ensures that citizens have access to the necessary tools, resources, and support systems to effectively generate, manage, and utilise data. Public funding and provisions facilitating private support (e.g. tax breaks or monetary incentives) are instrumental.

- **Financial resources** include sufficient funding to support the purchase of equipment, maintenance, and operational costs of CGD projects, and to fund training for CGD knowers. Availability of grants and sponsorships from governments, nongovernmental organisations, or private entities would contribute to kick-start projects but also to sustain long-term CGD initiatives and ensure that data remain available and maintained over time. States could retrieve the necessary monetary resources to support CGD capacity-building and initiatives by taxing revenues of data service streams (e.g. data centres).
- **Technological infrastructure** is a sine-qua-non condition of existence and survival for most CGD initiatives. This would include the availability of reliable devices and sensors that citizens can use to collect data (e.g. air quality monitors, water testing kits), access to high-speed internet and digital platforms that enable the transmission and sharing of data, and to a robust system for storing large volumes of data and processing capabilities to analyse the data collected. In this context, it would be important to consider the new data access rights provided by the Data Act to users of connected devices and how this can affect data collection practices in the context of CGD (see more about the opportunities and limitations of business to consumer (B2C) data sharing in Chapter 4 of this report).

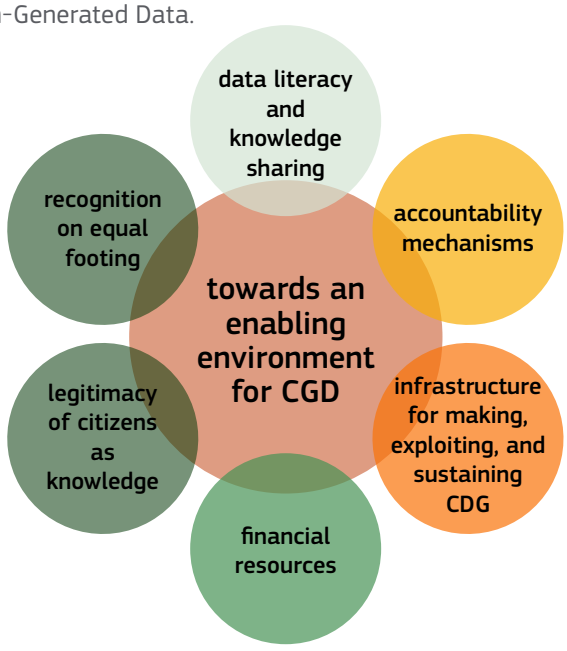
It has been observed that 'the degree and kind of empowerment environmental surveillance supports is determined by the manner in which surveillance data is made meaningful' (Ottinger, 2010, p.221, emphasis added). Thus, the third class of elements for an enabling ecosystem requires **pushing the boundaries of the regulatory framework** to make CGD meaningful.

- **Recognition on an equal footing**. Institutions often downplay the role of citizens and laypersons as knowledge creators. However, insights from multistakeholder governance, where all stakeholders are entitled to participate in policy decisions on an equal footing (cf., Raymond and DeNardis, 2015), can help attribute to CGD the role such date deserves. While disparities in resources and expertise can significantly undermine these ambitions, the institutional set-up would be geared towards attributing CGD the voice it deserves. This recognition would be tied to the development of adequate regulatory standards as a 'boundary-bridging' tool (Ottinger, 2010) for CGD to be recognised as valid sources of information.
- Within CGD, the **legitimacy of citizens as knowers** is fragile and contingent, drawing from various sources such as representation, experience, and even moral authority. It cannot be established in advance or permanently. However, a robust regulatory framework would significantly help in legitimising non-credentialed forms of knowledge. This aligns with the concept of 'supplementary democracy', which supports enhancing existing mechanisms to promote broader civic participation and greater accountability (van Rooy, 2004, p.137). Establishing the 'right to contribute information and have that information considered by appointed institutions' (Berti Suman et al., 2023, p.4) would further legitimise CGD.

Figure 5 depicts the six elements of an enabling ecosystem in relation to roles and responsibilities within the GDDS. It is to be read clockwise. Data literacy and knowledge sharing, in orange, are at the root of CGD, and are the joint responsibility of the CGD community and institutional actors. Accountability (in yellow) rests solely with CGD communities, while technological infrastructure, financial resources, legitimacy, and recognition on equal footing (in shades of green to indicate their incremental effect on CGD initiatives) are primarily the duty of institutional actors.

FIGURE 5.

Enabling ecosystem for Citizen-Generated Data.



Source: own elaboration.

6.4 Conclusions and policy recommendations

Citizen-generated data are non-traditional data sources that are created or collected, analysed, reused, and/or disseminated by individuals and groups outside the remit of the State and the industry, with or without the collaboration of specialists such as scientists and either not-for-profit or in combination with for-profit activities. They are at the core of a demand-driven approach to data production within the GDDS. The lifecycle of CGD includes six stages, divided in 'making' and 'exploiting' data. Citizens as knowers can either collect data that has been passively or actively generated or produce data from scratch; data, in both analogue and digital format, is then analysed. Once generated and analysed, citizens can exploit CGD in three main ways. They can promote environmental awareness and a just, sustainable transition by utilising data for informing policymaking, enforcing law or to raise awareness ('deploying data'). They can use data for co-designing legislation or to develop solutions, including products ('accelerating data'). Finally, they can use data for self-empowerment or to promote virtuous behaviour by others ('activating data'). These stages identify potential transformations implemented by means of data but are neither required nor mutually exclusive.

CGD encounter six main challenges to effectively contribute to environmental governance: data quality, data management, data ownership, sustaining participation, vulnerability of CGD communities, and governance and policy integration. An enabling ecosystem for CGD implementing the principle of careful knowing, centring the needs of marginalised knowers, empowering them and caring for them, would address these challenges by implementing community measures (literacy and responsibility), mobilising resources and infrastructure to support the grassroots generation and mobilisation of knowledge, and adapting the regulatory framework to meet the needs of CGD knowers.

As a result of the analysis, we can identify **three policy recommendations to foster an enabling ecosystem for CGD inspired to careful knowing and data justice**. The objective is to integrate, enable and give legitimacy to CGD, exploiting their potential in the promotion of environmental awareness and a just, sustainable transition.

i. Craft an incentive scheme to foster, support and sustain CGD initiatives redirecting revenues from the data economy (e.g. data centres, tech companies) to grassroots initiatives, including small and medium-sized enterprises (SMEs). This incentive scheme should consider both financial resources and technological infrastructure.

ii. In cooperation with relevant stakeholders, including RDAOs, develop adequate regulatory standards for CGD to be recognised as legitimate sources of information in decision-making processes. Revise these standards on a five-year basis to make sure they reflect the evolution of technology as well as policy problems.

iii. Promote CGD awareness regarding their potential as well as the role of RDAOs by means of EU-wide information campaigns and educational resources, in collaboration with public institutions (e.g. public libraries, the school system) and private entities (e.g. technology companies, the media system).