frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 02 April 2024
DOI 10.3389/fenvs.2024.1369001

OPEN ACCESS

EDITED BY
Claudia Schuetze,
Helmholtz Association of German Research Centres (HZ), Germany

REVIEWED BY
Uta Koedel,
Helmholtz Association of German Research Centres (HZ), Germany
Ayyoob Sharifi,
Hiroshima University, Japan

*CORRESPONDENCE
Koen Borghys,
koen.borghys@vub.be

RECEIVED 11 January 2024
ACCEPTED 11 March 2024
PUBLISHED 02 April 2024

CITATION
Borghys K, Vandercruysse L, Veeckman C, Temmerman L and Heyman R (2024), Localizing the sustainable development goals in smart and sustainable cities: how can citizen-generated data support the local monitoring of SDGs? A case study of the Brussels Capital Region. *Front. Environ. Sci.* 12:1369001. doi: 10.3389/fenvs.2024.1369001

Localizing the sustainable development goals in smart and sustainable cities: how can citizen-generated data support the local monitoring of SDGs? A case study of the Brussels Capital Region

Koen Borghys[1]*, Laurens Vandercruysse[2], Carina Veeckman[1], Laura Temmerman[1] and Rob Heyman[1]

[1]SMIT, Studies in Media, Innovation and Technology, Faculty of Social Sciences and Solvay Business School, Vrije Universiteit Brussel, Brussels, Belgium, [2]Department of Applied Economics, Faculty of Social Sciences and Solvay Business School, Vrije Universiteit Brussel, Brussels, Belgium

Introduction: The Sustainable Development Goals (SDGs) serve as the global reference framework for sustainable development endeavors. However, traditional data sources, including official statistics, fall short in effectively measuring SDG performance, due to substantial gaps in the availability of reliable, timely, actionable, disaggregated, and accessible information for policy formulation. This research explores the SDG monitoring potential of citizen-generated data to enhance local environmental in the Brussels Capital Region.

Methods: Employing a qualitative approach, the study first defines and maps essential characteristics of citizen-generated data for inclusion in environmental SDG monitoring. Subsequently, expert interviews refine these characteristics and explore design requirements tailored to the Brussels Capital Region.

Results: The research culminates in a framework linking essential citizen-generated data characteristics to design requirements, ensuring data suitability for local environmental SDG monitoring.

Discussion: This framework advances the existing literature by specifically addressing local environmental SDG monitoring through citizen-generated data. It offers practical insights for local stakeholders, particularly policymakers, aiming to overcome barriers to the uptake of citizen-generated data and ultimately enhances environmental SDG monitoring in the Brussels Capital Region. The framework's applicability in other regions or for non-environmental SDG indicators remains a potential avenue for future research.

KEYWORDS

smart & sustainable cities, data-driven policymaking, non-traditional data, sustainable development goals, environmental SDG monitoring, indicators, expert interviews, citizen science data

1 Introduction

The SDGs provide a global framework with a narrative to be implemented at the local level. Despite the inclusion of input from representatives of local governments during the formulation process, the SDGs remains a global agenda with an implicit focus on the national level. Consisting of a policy framework endorsed by all 193 member states of the United Nations (UN) and are set to be achieved by 2030. Currently, there is an ongoing process of localization facilitated by subnational actors, i.e., regional, and local entities (Greene and Meixell, 2017; Oosterhof, 2018).

The SDGs are associated with a large set of 248 indicators, which are intended to provide a comprehensive overview of progress while offering insights for future action. However, the reporting of these indicators is predominantly done at the national level, creating a challenge for subnational governments in identifying meaningful data sources to effectively measure the progress (Greene and Meixell, 2017; Oosterhof, 2018).

There is a critical need for high-quality, accessible, timely, reliable, and disaggregated data to facilitate the measurement of progress towards the SDGs (United Nations, 2020). Traditional data sources, such as official statistics, are proving inadequate for SDG measurement, necessitating innovative approaches for data collection (United Nations, 2016). At the same time, a "data revolution" is unfolding, driven by technological advances that have exponentially increased the volume and variety of available data (IEAG, 2014; Lämmerhirt et al., 2016; Fritz et al., 2019). The digital revolution, characterized by the widespread accessibility of broadband internet, mobile tools (e.g., dedicated apps and smart devices), and big data analytics, has transformed the landscape of information sharing, collection, and processing.

In addition, the rise of social media facilitates the promotion and encourages public participation in science projects at a range of scales, from local to global (De Rijck et al., 2020; van den Homberg and Susha, 2018; West and Pateman, 2017). This phenomenon opens unprecedented opportunities for societal transformation and environmental protection (Mahajan et al., 2022). Data generated directly by citizens to monitor, address, and instigate change on issues directly affecting them (referred to as citizen-generated data) holds the potential to offer a more precise and robust representation of progress. Often produced in real or near real-time, this data is firmly rooted in local contexts, amplifying citizen voices and perspectives on SDG progress, arguably including those typically marginalized and hard-to-reach populations. The production and utilization of citizen-generated data also fosters the direct, active and invested participation of individuals in advancing the SDGs (Higgins and Cornforth, 2015; Hecker et al., 2019; Moczek et al., 2021). While certain initiatives - such as eBird for conservation planning, and the European bird index for biodiversity and agicultural planning - have been successful in supporting environmental action, at both EU and Member State levels, the evidence points to a gap between the policy relevance and policy uptake (De Rijck et al., 2020).

In this research, we aim to offer nuanced insights to local governments regarding the barriers to the integration of citizen-generated data into monitoring initiatives like the SDGs and to suggest ways for optimization. To achieve this objective, we will construct a comprehensive framework delineating the data characteristics essential for suitable local SDG monitoring. Subsequently, we will translate these characteristics into requisite design requirements that must be considered to obtain these characteristics. The overarching goal is to formulate a set of generalized design principles that transcend diverse contexts. Citizen science projects can then refine and tailor these principles during the actual application phase.

The initial stage of this research involves desk research to identify pertinent data characteristics and their associated design requirements. Subsequently, the framework will undergo validation via a qualitative approach involving 10 semi-structured expert interviews (Van Audenhove and Donders, 2019). These interviews, conducted within the domain of citizen-generated data, will seek expert perspectives on the identified characteristics and design requirements.

The article's structure follows a logical sequence, commencing with an overview of current literature on citizen-generated data for local SDG monitoring. Subsequent sections detail the research methodologies employed and the data derived from the interviews. The article concludes with a summary and points for discussion, aiming to contribute to the discourse on citizen-generated data for local SDG monitoring.

2 Literature review

An expanding body of literature highlights the significance of localizing the SDGs. Although the SDGs were developed by and designed for national governments, the idea that there would be a crucial role for regional and local governments in the success of the goals was already present during the establishment of the Agenda 2030 (Reddy, 2016). Although the Agenda 2030 explicitly recognizes the key role of cities and municipalities by dedicating a specific SDG to Sustainable Cities and Communities (SDG 11), the importance of the local level within the SDGs goes further than SDG 11. The reasoning is that efforts to achieve certain subgoals will mainly have to be done in the context of local governments (cities and municipalities), which is the closest level of governance to citizens and is a necessary level to stimulate other actors to work towards these sustainability goals (UNDG, 2014). Various sources have estimated that around two-thirds or 65 per cent of the 169 SDG Targets will only be reached with a clear mandate and role for local (urban) actors in the implementation process (Cities Alliance, 2015; Lafortune et al., 2019; OECD, 2020).

These targets are linked to 62 per cent of all the Official Global Indicators (Ciambra et al., 2020), which brings us to the topic of local monitoring of the SDGs. To operationalise the SDGs, a form of progress measurement towards the goals will be necessary. However, the existing efforts done on the global (international) and national scale will not be sufficient at the local level because of various reasons. For example, national averages of scores on indicators can misinterpret realities on the ground and mask large regional or local disparities (SDSN & IEEP, 2019).

Local SDG monitoring not only necessitates a substantial volume of data but also demands data of high quality, extensive coverage, frequent availability, and spatial disaggregation (Fritz et al., 2019). Presently, SDG monitoring heavily relies on conventional and official data predominantly sourced from national statistical offices, like household surveys or administrative registers, often referred to as traditional data sources. Although essential and valuable, these traditional data sources exhibit several limitations (Ballerini and Bergh, 2021). Primarily are technical constraints, involving costly and time-consuming collection processes, along with limited spatial variation and coverage which results in infrequent data collection cycles, substantial data gaps, and unrepresentative samples that systematically exclude marginalized populations, such as vulnerable minorities, and sensitive issues (Lämmerhirt et al., 2016). Other often-mentioned constraints are related to the risk of manipulation by public officials to suppress contentious information, downplay their public institutions' challenges or artificially enhance their performances, and inadequacy of traditional data sources for capturing contextual information and local knowledge (Ballerini and Bergh, 2021).

Given the inherent limitations in traditional data sources, there is a growing consensus for SDG monitoring to complement them with unofficial and alternative data sources, often referred to as non-traditional data sources (Fritz et al., 2019; San Llorente Capdevila et al., 2020; Ballerini and Bergh, 2021). There are various examples of non-traditional data sources. Fritz et al. (2019) divided them into 5 main categories: Earth Observations, spatial data infrastructure, official sensor networks, commercial data and citizen-generated data. This last category consitutes data voluntarily generated and gathered by individuals, herein referred to as "citizens," with the purpose to monitor, advocate for, or instigate change in matters directly affecting them (Lämmerhirt et al., 2016; Ballerini and Bergh, 2021). It includes a wide variety of approaches and methods, with citizens utilizing a diverse array of technologies and participatory methodologies such as community-based monitoring, crowdsourcing on online platforms, or digital sensors (Lämmerhirt et al., 2016; Ballerini and Bergh, 2021). Typically, the production of citizen-generated data is initiated by citizens or civil society organizations and overseen by various intermediary entities, including non-governmental organizations (NGOs), academic researchers, private companies, and government agencies (Lämmerhirt et al., 2016).

It is worth noting that the existing literature lacks a universally accepted definition of citizen-generated data. Depending on the context, the term is frequently used interchangeably with concepts like citizen science data, community-driven data or participatory data (Lämmerhirt et al., 2016; Fritz et al., 2019; Ballerini and Bergh, 2021). Nonetheless, aligning with the perspective of Fritz et al. (2019), we position citizen science data as an integral part of the broader concept of citizen-generated data, where individuals actively contribute data as part of a scientific research process (McKinley et al., 2017; Serrano et al., 2018; Phillips et al., 2021). Presently, the prevailing focus within the literature predominantly centres on elucidating the significance of citizen science in contributing to the SDGs, as well as the utility of citizen-science data for monitoring SDG progress, as there are many examples of citizen science projects covering a diversity of domains that can contribute to the SDGs (Fritz et al., 2019; Phillips et al., 2021). We will thus use citizen science and citizen science data as umbrella terms that cover all these terms and diverse activities.

An examination of the potential role of citizen science in defining, monitoring and implementing the SDGs was initially presented in a discussion brief published by the Stockholm Environment Institute (West and Pateman, 2017; Müller et al., 2023) and since then the discourse on the impact of citizen science has evolved unto a growing research domain with growing output (Müller et al., 2023). Current research predominantly operates on the premise that citizen science can significantly contribute to the SDGs by facilitating the envisioned environmental and social transformations outlined in the SDG agenda, driven by its democratic principles (Sauermann et al., 2020; Alarcon Ferrari et al., 2021). These principles not only align with the SDG objective of "Leaving No One Behind" but also aim to enhance scientific productivity by involving society in the research process. Building upon these theoretical foundations, diverse stands of scientific literature have emerged. For instance, researchers have investigated how citizen science project coordinators evaluate alignment with the SDGs and the challenges encountered in contributing to them (Moczek et al., 2021). Moreover, studies have documented the actual and potential scope of various forms of contribution at project, local, national, and global levels (Fritz et al., 2019; Fraisl et al., 2020; Schleicher and Schmidt, 2020).

Despite acknowledging the potential of citizen science data for monitoring SDGs, the literature also highlights challenges and the limited integration of such data into monitoring and policymaking processes (De Rijck et al., 2020; Alarcon Ferrari et al., 2021). This aligns with broader challenges like data ownership (MacFeely, 2019), and the tradeoffs and negative impacts associated with big data and smart technology solutions for SDGs (Sharifi et al., 2024). For instance, financial limitations in implementing these technologies in cities can hinder progress towards specific goals like poverty reduction (SDG 1), zero hunger (SDG 2), and affordable and clean energy (SDG 7). Additionally Sharifi et al. (2024), identify trade-offs related to privacy, cybersecurity, infrastructure costs, biased decision-making that can exacerbate social inequalities (SDG 10) and limit accessibility due to the digital divide (SDG 9).

However, it is crucial to acknowledge that smart technologies also offer opportunities for achieving SDGs Sharifi et al. (2024). Highlight successful cases where responsible data governance practices and community-driven data collection (potentially linked to citizen science) helped mitigate challenges and contributed to positive outcomes. For example, initiatives that prioritize data privacy and security while empowering communities to collect and own their data can address concerns about data ownership and promote inclusive participation (SDG 10). Therefore, a nuanced approach that carefully considers both the potential benefits and drawbacks of smart technologies is necessary to ensure they contribute to achieving SDGs in a responsible and equitable manner.

This article extends the focus to understand how citizen science initiatives can be more effectively incorporated into local environmental SDG monitoring. Therefore, we delve into the conditions for enhancing the uptake of citizen science data into local governmental SDG monitoring mechanisms.

The focus centers on SDGs with a primary environmental emphasis, including clean water and sanitation (SDG 6), sustainable cities and communities (SDG 11), responsible consumption and production (SDG 12), climate action (SDG

TABLE 1 Value of citizen science data for SDG monitoring: Explanation of different features.

Dimension	Feature	Description	Value for the SDGs
Spatial	Spatial reference	Degree of inclusion of reference information	Location information can contribute to spatially explicit indicators
	Spatial resolution	Density and abundance of observations	Denser coverage than traditional surveys
	Spatial extent	Degree of (wide) geographical coverage and coverage of remote locations	Wide geographical coverage and remote locations
Temporal	Duration	Length of time data is collected	Regular or continuous data collection is well suited to the SDGs
	Resolution	Frequency of data collection	More frequent update cycles could fill temporal gaps in SDG indicators
Thematic	Coverage (subject)	Diversity of domains covered relevant for the SDGs	Multiple domains relevant to a range of SDG indicators
	Definition	The richness of the vocabulary that describes a particular subject area of concept	Richer vocabularies could fill data gaps (especially Tier II and III)
	Resolution	Level of detail provided by the vocabulary describing the data	More detailed vocabularies could fill data gaps (especially Tier II and III)
Process	Purpose	Implicit or explicit link with the SDGs	Potential implicit use of Citizen science data for SDG indicators
	Data processing	Cleaning, analyzing, and potentially modelling the collected data to make it usable for SDG indicators	Data processing can be aligned with indicator needs
	Data collection	Way in which data is collected (using a strict sampling design or opportunistic)	Data collection can be aligned with indicator needs
	Cognitive attention	Level of intervention needed during data collection (active vs. passive)	Both active and passive data sources can be valuable for SDG indicators
	Driver	Defines who is the driver of the proces (scientists vs. volunteers/ communities)	Both approaches can offer value for SDGs depending on the data needs
Data management (FAIR)	Findable	Data should be easy to discover both for humans and machines	New resources can be found for Tier II and III indicators
	Accessible	Mechanisms should be in place to allow access to the data	Persistence and openly accessible
	Interoperable	The data and its metadata should be understandable and usable by different software systems, allowing seamless integration	Integration with traditional data sources and across initiatives
	Reusable	Data is presented in a way that allows for reuse by others	Licensing, metadata and conforms to standards

Indicators are rank by data readiness for SDG monitoring, with Tier I indicators boasting clear definitions and readily available data, while Tier III indicators lack standardized methods for data collection.

13), life below water (SDG 14), and life on land (SDG 15) (UNEP, 2021). Citizen science activities span various environmental policy areas related to SDGs, making existing initiatives particularly relevant for providing valuable data to address these goals (McKinley et al., 2017; Nascimento et al., 2018; Serrano et al., 2018; Phillips et al., 2021).

The exploration begins with the scholarly work of Fritz et al. (2019), which introduces a conceptual model with five dimensions—spatial, temporal, thematic, process, and data management—each featuring distinct attributes highlighting the intrinsic value of citizen science data for SDG monitoring (Fritz et al., 2019). It provides a structured way to analyze various aspects of citizen science data across key dimensions, helping to understand its strengths and limitations for specific SDG indicators.

It breaks down the evaluation of citizen science data into different dimensions, this comprehensive approach ensures a thorough assessment of diverse factors influencing data suitability for SDG monitoring. Each dimension highlights positive aspects (attributes) of citizen science data, such as denser spatial coverage, more frequent updates and diversity of domains covered. This emphasizes the potential contribution of citizen science beyond traditional data sources. The framework provides specific values or descriptions for each feature, offering a concise understanding of how citizen science data performs in each aspect. Table 1 further explains the different dimensions and attributes.

As the framework serves as a theoretical model, this research objective is to customize it towards policymakers and citizen science practitioners. The primary aim is to enhance its practical utility, thereby amplifying its impact and fostering broader integration of citizen science for SDG monitoring. Beyond discerning the inherent value of citizen science, we seek to illuminate the requisites for extracting actionable insights from the data. We are particularly keen on identifying potential challenges and critical considerations that could hinder realizing the full potential of citizen science for SDG monitoring.

3 Methodology

3.1 Geographical focus

The Brussels-Capital Region (BCR), situated at the heart of Belgium, serves as the focal point of our analysis. Encompassing a surface area of approximately 161.4 km^2,[1] with a population exceeding 1.2 million inhabitants, it stands as Belgium's sole metropolitan area and is recognized as the capital of the European Union. Administratively, the BCR comprises 19 municipalities, including the city of Brussels.

3.2 Monitoring of SDGs in the BCR

Presently, a comprehensive framework or tool specifically tailored to aid local actors in monitoring SDGs within the BCR is lacking. Some existing initiatives, such as indicators.be,[2] Donut.Brussels,[3] OECD a Territorial Approach to the SDGs,[4] and SDSN European Cities SDG Index,[5] offer data and indicators for the region as whole. However, these initiatives predominantly operate at the regional level, overlooking the unique composition of the 19 individual municipalities within the BCR.

At the local level, some municipalities in the BCR collect data related to themes within the environmental SDGs as part of own developed climate plans in response to a project call initiated by the regional government. These efforts mainly rely on traditional data sources, with the potential of citizen science data sources largely remaining untapped.

3.3 SDG in ACTION

The growing demand for support for local implementation of the SDGs resulted in the SDG in ACTION PhD project.[6] This applied PhD, funded by the Brussels Capital Region (Innoviris) and implemented by Studies in Media and Innovation (SMIT) and IDEA Consult. The primary objective is to provide support to local public and private stakeholders within the BCR in their efforts to monitor progress towards local SDG targets. This study is done in the context of this research project.

3.4 Citizen science projects in the BCR

The BCR today hosts several noteworthy citizen science data projects, particularly in the domain of Air Quality monitoring, including projects such as CureuzenAir BXL, compAIR, AIRCasting, expAIR. Despite the existence of these projects, there is currently no comprehensive public portal that consolidates information on past, ongoing or upcoming citizen science projects within the region.

However, there are available resources that offer partial insights into the landscape of citizen science in the BCR. The EU-citizen.science website[7] features some projects associated with the BCR, but the coverage is somewhat limited. Furthermore, the website from the Vrije Universiteit Brussel[8] provides an overview of some of the projects researchers from the university are involved with.

3.5 Data collection method

Our data collection adopts a qualitative methodology, employing semi-structured interviews with data experts in citizen science data. This method aims to extract bottom-up insights into various aspects related to the data generated during citizen science data projects.

In the initial phase, a comprehensive literature review and desk research were conducted to orient the research, identify relevant experts, and construct a topic list to guidance the interviews (Van Audenhove and Donders, 2019).

The interviews were conducted using a semi-structured topic list, providing flexibility to guide the conversation and adapt questions based on the interviewee's responses. Open-ended questions were employed to allow interviewees to provide detailed and nuanced responses, thereby enabling a deeper understanding of their perspectives and often yielding valuable additional information. The interviews were structured into five broad, predefined sections:

1. Welcome and introduction: Communicating the purpose of the interview, its relation to the SDG in ACTION project, and outlining the interview process.
2. Introductory questions and general knowledge: Exploring the interviewees' familiarity with the SDG framework and non-traditional data sources.

1 https://bisa.brussels/wist-je-dat/162-4-km2-is-de-oppervlakte-van-het-brussels-hoofdstedelijk-gewest

2 https://indicators.be/en/t/SDG/

3 https://donut.brussels/en/brussels/en-macro-portrait/

4 https://www.oecd-local-sdgs.org/index.html

5 https://euro-cities.sdgindex.org/#/

6 https://smit.vub.ac.be/phd-project-sdg-in-action-data-driven-coaching-of-local-communities-in-achieving-sdg-goals

7 https://eu-citizen.science/projects?keywords=&orderby=&country=BE&status=&hasTag=&difficultyLevel=&topic=&participationTask=

8 https://www.vub.be/en/our-research/our-vision-and-mission/participatory-and-community-based-research/citizen-science

TABLE 2 Background of participants and expertise.

Type of organization	SDGs	Field
Knowledge-based SME, research institute, environmental agency, international organization, museum, botanical garden	6, 11, 13, 15, general (environmental Citizen science projects)	Odor, air quality, biodiversity, mobility, water

TABLE 3 Adapted framework with overarching layers.

Overarching layer	Dimensions	Attributes
Data Scope	Spatial	Spatial reference, Spatial resolution, Spatial extent
	Temporal	Duration, Resolution
	Thematic	Coverage (subject), Definitions, Resolution
Data governance	Process	Purpose, Data processing, Data collection, Cognitive attention, Driver
	Data management (FAIR)	Findable, Accessible, Interoperable, Reusable
	Data quality	Evaluation with traditional measures or measures more tailored to citizen science data

3. Citizen-generated data sources and the SDGs: Delving further into different elements of the framework making abstraction of the local context in the BCR.
4. Citizen-generated data sources for the BCR: Addressing questions on the concrete applicability of citizen-generated data for the BCR and ideas for future projects.
5. Closing: Conduct a short debrief, providing information on the next steps in the data analysis process.

In the subsequent phase, participants were selected based on specific characteristics relevant to the research question. The selection involved reaching out to the Flemish Knowledge Centre for Citizen Science (Scivil)[9] to support the identification of suitable data experts. We employed a purposive sampling strategy, a non-probability method that involves selecting participants based on specific characteristics relevant to the research question. This strategy aimed to include a diverse range of experts and projects, obtaining varied perspectives and activities without seeking population generalization (Bryman, 2012).

Invitations were extended to experts with proficiency in citizen science data related to diverse environmental topics, including air and water quality monitoring, biodiversity, and with knowledge on a variety of data gathering techniques. The interviewed experts possess various backgrounds, representing both governmental and non-governmental stakeholders. Further details about the background of the participants and their expertise are presented in Table 2 below.

The expert interviews, totaling 10, were conducted through video conferencing on Teams, each lasting approximately 1.5 h. Given the multilingual nature of the BC, a multilingual approach was adopted, conducting interviews in English, French or Dutch, based on the participants' preferences. The interviews were systematically recorded, automatically transcribed, and supplemented with notes for subsequent data analysis where thematic analysis was employed to identify and analyze themes within the data (Herzog et al., 2019).

9 https://www.scivil.be

4 Findings

This study adopts the framework outlined by (Fritz et al., 2019) to structure the insights gathered from our expert interviews. However for improved clarity, we introduce two overarching layers that consolidate the original set of five dimensions. The Data scope layer encompasses the spatial, temporal, and thematic aspects of citizen science data. It aligns with the prevalent understanding of data scope as the breadth and depth of information captured within a dataset (Abraham et al., 2019). Conversely, data governance focusses on the management, processing, and overall governance structure of the citizen-science projects. It incorporates the formalization of data policies, standards, and procedures (Abraham et al., 2019), encompassing the data management and process dimensions from the original framework. This way it emphasizes the interconnected nature of data collection, processing and sharing. Additionally, we integrate data quality into the framework as a separate dimension under data governance, recognizing its importance as a barrier to citizen-generated data utilization in monitoring efforts (Fritz et al., 2019). This dual-layered approach provides a comprehensive structure for organizing and analyzing expert insights. It enhances the clarity and coherence of our analysis, facilitating a holistic understanding of the various facets associated with citizen science data within data scope and data governance. Table 3 below outlines the key elements of the adapted framework, including layers, dimensions, and attributes.

4.1 Data scope

This layer plays a crucial role within the overall framework and delves deeper into three key dimensions that define the scope and suitability of citizen science data for monitoring SDGs: 1) Data extending traditional data boundaries, 2) fit for purpose, and 3) data representativeness.

4.1.1 Data boundary extension

Our experts underscore the constraints of traditional data sources regarding the coverage, resolution and extent these sources offer for monitoring within their domains of expertise like biodiversity, air- and water quality. Extending beyond these conventional boundaries, citizen science data has the potential to transcend the limitations inherent in traditional data sources. This can take various forms, e.g., spatial, temporal, or thematic, leading to attaining a level of granularity that aligns with the specific information requirements of policy formulation. Moreover, using citizens' capacity to articulate their sentiments and preferences on various matters can also be a way they go beyond the boundaries of traditional data. This approach facilitates a nuanced understanding of issues like pollution, surpassing the confines of (annual) emission measurements by incorporating citizens' perceptions and experiences, for example, including the assessment of citizens' perceptions of odors.

Citizen involvement enables a cost-effective expansion of measurement points, supplementing traditional data sources and ensuring geographical coverage in areas that may otherwise be overlooked by the traditional governmental measurement stations like remote, hard-to-reach, or inaccessible locations. The "CurieuzenAir BXL" project in the BCR, serves as a good example, as it employed citizen science sensors for air quality monitoring (concentration of NO2) in the region to provide a more intricate perspective on air quality. By involving citizens asking them to attach a sensor to their houses it covered Air Quality at street level.

Telraam serves as another illustration, employing sensors affixed to citizens' windows to provide both citizens and local administration with a comprehensive portrayal of neighborhood traffic. The insights derived from such initiatives surpass the granularity achieved through a limited number of counting loops or cameras installed by a local government.

4.1.2 Fit for purpose

While the expansion of existing traditional data boundaries is a prerequisite for citizen science data to become interesting for governments to use the data in monitoring efforts, our experts emphasize the need for a careful balance to ensure that the collected data aligns proportionally with these monitoring objectives. The gathered data should be proportional and adequate for the intended monitoring purposes, which could be in conflict with typical citizen science approaches where the purpose is defined before data collection begins. In conventional citizen science projects, clear goals and objectives are established, and data is collected specifically to address those questions. This approach ensures that the data is relevant, efficient, and manageable for analysis and decision-making in the context of the project.

Nevertheless, there are instances where citizen science projects may adopt a more open-ended exploration approach, allowing participants to generate valuable data without predefined objectives. While this may initially seem divergent from the traditional method, such exploratory projects can yield data that, under retrospective assessment, proves valuable for specific monitoring efforts, particularly at the local level.

This highlights challenges arising from the abundance of citizen science data for local governments. While such data offers enhanced granularity, it simultaneously presents hurdles—particularly in terms of increased data processing demands. Local governments, constrained by human and technical capacity and expertise, may find managing this influx challenging. Additionally, for monitoring purposes, data should offer information that is easily interpretable, necessitating only the granularity essential for its intended purpose and thus appropriate scales.

The pursuit of super-accurate data may not always be necessary (Bowser et al., 2020). In certain scenarios, a reduced level of accuracy may still be sufficient for monitoring purposes. An often-cited example by our experts is related to the temporal scale. For monitoring purposes, real-time data is often not an absolute requirement. Rather, the data should be updated at a frequency corresponding with the analysis or monitoring needs of the government.

Furthermore, a call is made for broader Application Programming Interface (API) capabilities for specific policy applications and operational decision-making. APIs function as intermediaries that enable seamless communication between distinct software programs, thereby facilitating real-time data visualization on dashboards and remote adjustments to sensor settings. An illustrative example involves the case of Telraam, where adjusting the API's data measurement frequency enabled a more accurate assessment, showcasing the need for tailored API solutions to optimize data usability for policy applications. In the context of a school street,[10] operational from 7:30 to 8:30 am and 3:30 to 4:30 pm, the sensors initially recorded data on an hourly basis, spanning from 7:00 to 8:00, 8:00 to 9:00 am, and so forth. Consequently, the data encompassed half-hour intervals, during which the school street was both opened and closed. To address this limitation, a modification to the API was implemented, shifting the data measurement frequency to a quarter-hour basis. This adjustment proved instrumental in comprehensively capturing and mapping the impact of the school street on the surrounding environment.

To allow for adequacy in terms of available data, citizen science projects need to profoundly understand the data needs of the local government for monitoring purposes. In the context of exploring viable data for SDG monitoring, this necessitates a thorough examination of the SDG framework itself. While ensuring thematic alignment with the SDGs is crucial, our experts express the challenge regarding the limited knowledge among citizen science projects of the SDG framework and the difficulty of understanding its intricate nature.

Concerning the first element, our experts highlight a limited integration of the SDGs within citizen science data projects. Instances where the SDGs are incorporated tend to be associated with mandatory links to proposals for external funding, indicating a somewhat instrumental use rather than intrinsic alignment. Addressing the second element, our expert's express challenges

10 A public road in the vicinity of an educational establishment where motor vehicles are temporarily barred at the entrances during certain hours.

arising from the complex nature of the SDGs. One expert mentioning the 'academic nature' of the SDGs as a barrier, presenting a difficulty in comprehending the framework and identifying relevant indicators. Another expert, focusing on citizen science data related to biodiversity, exemplifies the struggle in monitoring specific SDG targets (15.2 and 15.9) without having a comprehensive understanding of the broader SDG framework. These examples underscore the common difficulties faced by projects, emphasizing the need for a more accessible interpretation of the SDGs. Existing knowledge often centers around the 17 goals themselves, with limited awareness of the underlying framework of subgoals and indicators. There is a call for translating the SDGs into a format adapted to the local context of the BCR, ensuring a clearer and more relevant perspective aligned with local needs.

4.1.3 Data representativeness

A recurrent theme highlighted by our experts revolves around the vital consideration of the representativeness of the data being collected by citizen science data projects. Inclusivity emerges as central, emphasizing the need to avoid data biases and ensure representation in line with the fundamental principle within the SDG framework of "Leave No One Behind" (LNOB) (Lämmerhirt et al., 2016). The literature underscores citizen science data as promising in this regard, enabling the capture of bottom-up insights on under-reported issues, facilitating the expression of counter-narratives of sustainable development, and the genuine empowerment of all voices (Lämmerhirt et al., 2016).

While citizen science data offers promise in terms of inclusivity, current trends often involve participation biased towards individuals with higher socio-economic status. As articulated by one expert, "challenges arise in achieving diversity, with prevailing trends often featuring the participation of economically privileged individuals, typically of white ethnicity, male, possessing advanced degrees, and of an older age." Consequently, projects must broaden their participation pool beyond those who already possess resources like time and capital to engage in citizen science activities.

It is important to recognize that citizens engage in citizen science activities for various reasons, ranging from personal interest in the topic, financial incentives for data acquisition, a desire for knowledge enrichment throughout the project (such as understanding the functioning of sensors or the overarching topic), a sense of community belonging, or a genuine intention to contribute positively. Acknowledging and catering to these diverse motivations is crucial for attracting a wide range of participants.

Successful citizen science projects should actively involve underrepresented groups, fulfilling the overarching purpose of democratizing knowledge for policy formulation and implementation. To address this challenge, deliberate efforts are needed to include these underrepresented groups and amplify unheard voices.

The importance of representativeness varies depending on the type of data being collected, as the imperative for representativeness varies across domains. In biodiversity topics, such as counting vulnerable species, the characteristics of the citizen collecting the data may have less impact. However, in areas like environmental issues linked to socioeconomic factors, representativeness becomes critical. This is exemplified by the results of CurieuzenAir BXL project, were 3,000 participants across diverse socioeconomic areas within the territory of the BCR used citizen science sensors to map air quality, revealing correlations between air quality and the socio-economic status of specific neighborhoods.

Recognizing that not all projects may have the means to collect data on such a broad scale, our experts suggested measures to enhance the representativeness of the data sample, by collaborating with intermediaries. One suggested approach is to engage with social organizations, directly involving them to facilitate the deployment of citizen-science sensors. However, it is essential to acknowledge the financial constraints often faced by these organizations. Another viable method involves building partnerships with organizations to access their knowledge regarding underrepresented groups. For example, schools can be valuable contributors, leveraging their existing data to delineate the proportion of certain demographics within a population, thereby augmenting representativeness. In the context of the BCR, Flemish schools participating in initiatives like the Flemish GOK-Indicators[11] related to equal educational opportunities, can provide data on parameters related to individuals in need of support.

Privacy concerns, in compliance with regulations like the GDPR, demand careful attention. Issues may arise in accessing vulnerable groups, such as the Roma community, highlighting the ethical and practical challenges of ensuring inclusivity in citizen science initiatives.

4.2 Data governance

Within the overarching framework, data governance plays a critical role. It is important to understand its two key components, data process and data management.

4.2.1 Data process

4.2.1.1 Citizen science for the long haul

For data to be truly valuable, it must extend beyond the temporal boundaries of individual projects. Our experts highlighted the pronounced challenge of maintaining projects and ensuring continuity in data gathering. Particularly given the project-based nature of many citizen science initiatives, it requires attention to the intermittent nature of data collection inherent in such projects. Seamless integration of these projects into governmental monitoring initiatives necessitates regular campaigns or continuous data collection.

This often originates from limited funding, restricting projects to data collection only over a limited period. The financial aspect is a significant consideration in the realm of citizen science projects. For instance, while citizen science sensors may offer cost-effective solutions in specific domains, it is imperative to acknowledge that maintaining projects over an extended period can still incur substantial expenses. A case in point is the CurieuzenAir BXL project, where the tubes used to measure air quality incur regular laboratory costs of around 10€. Although manageable over a short period, these costs accumulate

11 *Gelijke OnderwijsKansen* or Equal Education Opportunities. Indicators on language spoken at home, school allowance, educational level parents.

over time. Thus, to enhance the sustainability of citizen science initiatives, a comprehensive understanding of the potential long-term expenses is crucial. These long-term costs are related to project maintenance, including for example, data collection, data management, sensor calibration, data storage and hosting,[12] and ongoing support. A concrete illustration of such costs is evident in the case of one of our experts, who incurs a monthly expenditure of € 1,000 for data hosting through Amazon Cloud Services.

Conversely, certain factors contribute to cost reduction, such as leveraging voluntary technicians for the maintenance of measurement instruments like sensors. For instance, citizens actively engage in cleaning water quality sensors and conduct checks by capturing and sharing pictures with a centralized database. If organizers identify trends in the sensor data that appear inaccurate, they can remotely validate the sensor quality. This collaborative effort not only reduces costs but also enhances the efficiency of data maintenance.

Furthermore, the maturity level of technology emerges as a significant factor influencing costs. Advancements in technology can lead to more cost-effective solutions, streamlining the overall expenses associated with citizen science initiatives. An expert specialized in citizen science water sensors elucidated that the maturity level of sensor technologies varies across different fields. Notably, air quality sensors have advanced significantly, providing cost-effective options that facilitate straightforward installations by citizens, thus being considered "plug and play." This high level of maturity in the technology translates into lower usage costs. In contrast, water quality sensors may lag in maturity with the development and accessibility of affordable and user-friendly water quality sensors still in nascent stages, making them less seamlessly integrable or "plug and play" compared to their air quality counterparts, potentially contributing to higher costs.

Some initiatives have transitioned into long-term data providers, establishing a business model to operate as sustainable service providers for local governments. Telraam, active in the BCR, is a prime example. Originating as a project that developed a traffic-counting sensor, it transformed into an enterprise with its own business model, serving as a sustainable data provider to local governments. The success of such initiatives relies on demonstrating the added value of their data to the local government, complementing existing traditional data sources fulfilling specific data needs, and attaining the requisite volume to maintain cost-effectiveness.

Identifying and addressing data blind spots regarding traffic counting, particularly in less-travelled secondary roads, holds significance for local administrations, prompting a willingness to invest in such data. Traffic count data proves valuable not only for its direct application, but also as proxy data for urban livability, encompassing factors like traffic density and share per type of road users, or air quality assessment, particularly in understanding emissions related to motorized traffic.

12 Storing the data on a stable and accessible web platform (https://www.gbif.org/data-hosting); costs can incur costs for a subscription on a cloud-based platform.

Notably, the design of the sensor itself contributes significantly to the success of initiatives like Telraam. Tailoring sensors for user-friendliness and application specificity, enhances citizen engagement, as evidenced by Telraam's approach. This underscores the importance of citizens recognizing the value of the data they collect, thereby increasing their likelihood of remaining long-term engaged data providers. Again, Telraam serves as an exemplary model in this context. For citizens, the data holds significance as it enables them to objectively address concerns about local traffic conditions in their streets, fostering a more effective dialogue with their local government. This not only contributes to citizen empowerment but also cultivates a heightened sense of ownership and involvement in community affairs.

Intrinsic motivation also emerges as a vital factor for the sustainability of citizen science projects, where managing engagement fatigue poses a specific challenge. Sustaining motivation for data sharing and continued participation beyond the project lifecycle requires careful expectation management and clear communication about long-term benefits. Strategies to actively engage and motivate participants in the data collection process are imperative. While organizers may influence certain elements, such as sensor design and engagement strategies, external factors like citizens moving to different locations may remain beyond their control. Open communication channels and a transparent process help build trust and encourage active participation. This also is demonstrated in the need for legal and ethical considerations as discussed below.

4.2.1.2 Ethical and legal considerations

Our experts underscore the significance of ethical and legal considerations, particularly in interactions with participating citizens during projects. Citizens should be well-informed about the handling of their data: Transparency and clarity regarding the use and transfer of data are crucial to establishing trust. Consequently, clear agreements, such as an informed consent form, should be in place, outlining the terms and conditions governing the project's objectives, the purpose of data collection, data usage, ownership and potential data transfer to (governmental) data repositories.

4.2.2 Data management

4.2.2.1 FAIR principles in practice

In the sphere of citizen science initiatives, effective data management plays a pivotal role, encompassing a variety of tools dedicated to the facets of data collection, storage, and dissemination. Its significance lies in facilitating collaboration between citizens and institutions throughout the entire project lifecycle. Incorporating FAIR principles—ensuring data is Findable, Accessible, Interoperable, and Reusable—form a strong foundation for data management. It ensures high quality data that can be effectively utilized for scientific research, policymaking, and community engagement (Wilkinson et al., 2016; Bowser et al., 2020; San Llorente Capdevila et al., 2020; Fritzenkötter et al., 2022).

Exemplary projects, such as WorldFAIR in the realm of Biodiversity, are cited by our experts as model for implementing FAIR principles and advancing data management practices in citizen science initiatives. These endeavors contribute to the creation of a robust and accessible data ecosystem beneficial to both the scientific community and the broader public. However, despite the significant improvements FAIR principles offer, specific concerns remain, including data quality, software compatibility, and content related trustworthiness (Koedel et al., 2022). Addressing these challenges necessitates additional approaches for ensuring effective data linkages and joint interpretation. While our experts recognize these challenges, they also offer valuable insights into key elements supporting the implementation of FAIR principles.

Notably the publication of data was highlighted, emphasizing the importance of publishing citizen science data in a professional manner, in adherence to appropriate metadata standards. While certain fields lack metadata standards, the development of such standards becomes essential to establish a structured framework for describing data and its origin. This contributes to enhanced data interoperability, facilitating seamless collaboration and analysis across diverse projects and domains. For instance, the use of the INSPIRE Metadata Regulation, a common metadata standard for the Infrastructure for Spatial Information in the European Community (INSPIRE),[13] ensures transparency in data provenance - clarifying the origin, creation, and propagation of the data (Imran and Agrawal, 2017). Standardizing data in a uniform and interoperable manner is also exemplified by the Darwincore in Biodiversity data which includes a glossary of terms intended to facilitate the sharing of information about biological diversity by providing identifiers, labels, and definitions.[14]

Our experts furthermore underscored the importance of making data accessible in an open data format—preferably a machine-readable format - as still in many cases data is made available in human-readable format (excel, pdf) (e.g., CSV, JSON, XML) (EC JRC, 2020). Particularly noteworthy is the progress made in the biodiversity data domain through the implementation Linked Open Data (LOD). Notably, certain other fields have already embraced LOD practices, showcasing it as a promising model for adoption in diverse citizen science initiatives. This trend suggests that leveraging linked open data principles can contribute significantly to enhancing the accessibility and interoperability of data in various domains within the realm of citizen science.

The integration of data from citizens-science data projects and applications into monitoring systems presents challenges, particularly in aligning citizen science projects with existing proprietary applications of local governments. This is related to the data structure, but also to the way the data is introduced into these governmental systems. While the proposition of constructing an appropriate API for integrating data into monitoring systems is widely acknowledged (EC JRC, 2020), one of our experts mentions that challenges arise due to several aspects, like the ownership of data by citizens, or when the data enters the governmental system, the control is lost over the data from the report or the feedback back to the citizens which is supposed to be given.

4.2.2.2 Data quality assurance

The recurring theme of concern surrounding data quality is underscored by our experts, who recognize factors contributing to lower data quality, such as sensor quality and external influences like power cuts. Proposed solutions include efforts to improve the quality of low-quality sensors and employ techniques like time series analysis to address outliers. Additionally, the use of expert calibration algorithms has been suggested.

To preserve data quality in certain fields, a multi-layered approach to data sourcing is showcased, involving the integration of data sets from distinct sources, with the inclusion of citizen science sensor data. In the context of air quality measurements, three layers are delineated. The first layer consists of reference stations adhering to European norms, thereby upholding high standards of testing and calibration, ensuring reliability and accuracy. The second layer encompasses commercial sensors, which, despite their cost-effectiveness, adhere to rigorous testing and calibration criteria. The third layer are citizen science sensors, characterized by lower components and minimal calibration, manifesting limitations in terms of sophistication and calibration precision.

As Table 4 below illustrates, each layer has its strengths and weaknesses. Reference stations adhere to the standard but lack widespread coverage. Commercial sensors offer good value but might miss local details. Citizen science sensors, while valuable for broader understanding, may require further refinement for accuracy. Notwithstanding potential limitations inherent in each layer, the integration of data from all three layers is crucial for a comprehensive and nuanced understanding of air quality, from large scale trade trends to localized hotspots.

Given the prevalence of low-cost sensors in citizen science, where calibration is often a one-time procedure at the factory, and discrepancies among sensors may persist even after calibration, a recommended practice is subjecting low-cost sensors to a sensor comparison experiment for validation prior to distribution to citizen.

The integration of APIs can significantly enhance the quality of citizen science data. APIs enable direct data storage within databases upon collection, streamlining the quality assurance and quality control (QA/QC) processes. They contribute to improved data integrity by identifying outliers, detecting drift, and ensure consistency with neighboring data points. Established data quality measures traditionally applied to official data, such as positional and thematic accuracy, temporal currency, completeness and representativeness across spatial and temporal dimensions, and appropriateness for the intended purpose (Fritz et al., 2019), can seamlessly be implemented within citizen science projects through the use of APIs.

Apart from actual data quality, a significant aspect revolves around potential misperceptions concerning data quality, recognizing that citizen science data may not always universally be accepted as qualitatively sound. Such considerations underscore the importance of end-users of the data, like civil servants, policy- and decision-makers, possessing expertise in data interpretation. It is emphasized in our interviews that a singular anomalous data point does not necessarily discredit the entirety of the dataset, aligning

13 https://rdamsc.bath.ac.uk/msc/m66

14 https://dwc.tdwg.org

TABLE 4 A comparison of three layers of Air Quality Sensors.

Layer	Description	Data accuracy	Spatial resolution	Temporal resolution	Strengths	Weaknesses	Example use cases
Reference Stations	High-end sensors adhering to strict European norms, with regular testing and calibration	Highly accurate and standardized	Low (regional, national)	Low (Hourly or lower) (days or months)	Highly reliable and accurate data	Expensive to deploy & maintain	Official reporting, regulatory purposes, long-term trend analysis
Commercial Sensors	Cost-effective alternative to reference stations, meeting established testing and calibration criteria	Moderate accuracy, may vary between models	Medium (urban areas, specific locations)	May vary: Medium to high (minutely to hourly)	Affordable option, wider coverage than reference stations	May have lower accuracy than reference stations	Air quality monitoring for specific areas, research studies
Citizen Science Sensors	Low-cost sensors deployed by individuals or groups, with minimal calibration	Lower accuracy, variable across deployments	High (street, neighborhood)	May vary: Medium to high (minutely to hourly)	Large-scale deployment possible, community engagement, hyperlocal data	Limited accuracy & precision, potential data quality issues	Community-based air quality monitoring, identifying pollution hotspots, research studies

with the concept of fitness for purpose and emphasizing the contextual significance of data quality and its application (Bowser et al., 2020). Furthermore, the depth and quality of the data processing also play a crucial role in shaping the final dataset and thus the information presented to the end-users.

The complexity of data analysis is linked to an understanding of legal thresholds, as exemplified in the context of air quality data. Recognition is essential that the presence of an average threshold does not necessarily deem the data unreliable if occasionally surpassed. In the case of air quality assessments, where various temporal thresholds are in place, a singular instance of surpassing the threshold does not inherently compromise the overall quality of the measurement data.

Expertise regarding data interpretation within local governments is paramount to ensure that collected data is effectively analyzed and interpreted to derive meaningful insights. Support from higher governments can play a vital role in providing the necessary resources, guidance, and validation for the data interpretation process. However, treating local governments as a homogeneous group does not correspond with reality, as various sizes, organizational capabilities, and data needs were highlighted during our interviews. For instance, the BCR encompasses 19 municipalities, with the city of Brussels standing out due to its significant population and hosting numerous EU-related and international institutions, in contrast to smaller municipalities within the BCR. A higher government within the BCR can play a pivotal role, not only in providing expertise regarding data interpretation but also regarding infrastructure for data sharing and analysis.

4.3 Refined framework

Building upon the foundational work of Fritz et al. (2019), this revised framework articulates essential data characteristics and design requirements for effective citizen science data integration into local environmental SDG monitoring. The framework seeks to provide a nuanced understanding of the intricacies involved, fostering collaboration between local governments and citizen science projects.

Under the data scope layer, crucial characteristics are identified to render citizen science data valuable for local SDG monitoring initiatives. These encompass:

1. Data extending traditional boundaries: Citizen science data should surpass limitations of traditional data sources, expanding spatial, temporal, and thematic coverage, including the incorporation of citizens' perceptions and experiences.
2. Fit for purpose: The data should be aligned with monitoring objectives, finding a balance between the right granularity and manageable data processing.
3. Representativeness: Emphasizes the importance of inclusivity to avoid biases in citizen science data.

Under the data governance layer, dimensions are perceived as constituting design requirements, outlining necessary processes and practices for well-designed citizen-generated data projects. These encompass:

1. Sustainability: Highlights the need to maintain data gathering beyond individual citizen science projects for continuous data collection.
2. Ethical and legal considerations: Emphasizes the importance of ethical and legal considerations in citizen science projects, encompassing transparency, informed consent, and considerations of data ownership.
3. Effective data management: Explores the practical implementation of FAIR principles (Findable, Accessible, Interoperable, Reusable) in citizen science data management. Emphasis is placed on professional data publication, adherence to metadata standards, and the adoption of open data formats.
4. Data quality assurance: Recognizes the importance of data quality in citizen science projects, addressing challenges including sensor quality, external influences, and the importance of the right data interpretation.

The revised framework furnishes a more intricate and nuanced comprehension of the various aspects influencing the value of

TABLE 5 Overview of the revised framework with data characteristics and design requirements.

Overarching layer	Dimension	Description	Challenges	Examples
Data scope (data characteristics)	Extending traditional boundaries	Data should go beyond traditional sources (spatial temporal, thematic, perceptions)	Data needs to align with monitoring purposes (see below)	CurieuzenAir BXL (Air quality), Telraam (traffic monitoring)
	Fit for purpose	Data aligns proportionally with monitoring objectives	Monitoring purposes vs. Project goals Abundance of data	API adjustments to optimize data usability (Telraam)
	Representativeness	Inclusive data collection to avoid biases	Engaging underrepresented groups, acknowledging diverse participants motivations, privacy concerns	Collaboration with intermediaries, building partnerships
Data governance – process and data management (design requirements)	Sustainability	Maintaining data gathering beyond individual projects	Limited funding, technology maturity, intrinsic motivation by participants	Transition towards a sustainable service provider (Telraam)
	Ethics & Legality	Importance of ethical and legal considerations	Transparency and clarity regarding use and transfer of data	Clear informed consent forms
	Effective data management: FAIR principles in practice	Effective data management facilitated by:	Integration of citizen science data into monitoring systems	WorldFAIR in biodiversity, LOD
		• Professional data publication		
		• Metadata standards		
		• Open data formats		
	Data quality assurance	Importance of data quality and its perception	Sensor limitations, external influences, lack of expertise in interpretation by end users	Multi-layered data sourcing (reference stations, citizen sensors), API usage

citizen-generated data for governmental monitoring. It addresses challenges, provides examples, and offers practical considerations across diverse aspects, positioning it as a valuable tool for comprehending and executing citizen science initiatives in the context of governmental monitoring. Table 5 below presents an overview of the revised framework.

5 Limitations

The research methodology employed in this article presents certain constraints that warrant consideration. The geographic focus is primarily on the BCR, and caution should be exercised when generalizing the findings to other regions. The intentional choice to closely examine an individual locality underscores the research emphasis on advocating for a bottom-up refinement of SDGs, aligning with existing research recommendations.

While the presented methodology offers a valuable framework, its application in diverse contexts necessitates careful consideration of site-specific factors such as the number of relevant citizen science initiatives providing data and the degree of local administrative support. These factors significantly influence its potential contribution to SDG monitoring.

A limitation arises from the sample set used for interviews, which primarily comprised of data experts and which was limited to 10 interviews. This selection choice imposes certain restrictions on the breadth of perspectives considered.

Furthermore, while the study engaged with data experts and subject specialists, it did not involve stakeholders from the local governments themselves, especially at the political level. Future research endeavors may benefit from exploring how insights from this perspective align with of diverge from those gathered, thereby contributing to a more comprehensive understanding. It is acknowledged that local (political) preferences regarding data-driven governance exist, and their inclusion in future investigations would offer valuable insights into these nuanced dynamics.

6 Conclusion—implications—applications

This article emphasizes the growing recognition of the crucial role of local governments in achieving the SDGs and thus the need for monitoring them at the local level. Traditional data sources, while essential, have limitations, leading to a call for supplementing them with non-traditional sources, like citizen science data. Existing literature explores the significance of citizen science data in contributing to and monitoring SDG progress, acknowledging challenges and limited integration into policymaking. It extends this discussion by focusing on how citizen science data can be effectively incorporated into local environmental SDG monitoring, investigating the barriers for uptake of citizen science data into local SDG monitoring efforts, with a focus on environmentally relevant SDGs. The exploration builds upon a conceptual model introduced by Fritz et al. (2019), which outlines dimensions highlighting the intrinsic value of citizen science data for SDG monitoring.

The study focusses on the BCR and involves qualitative data collection through 10 in-depth semi-structured interviews with citizen science data experts from various backgrounds. We started restructuring the original five dimensions of the conceptual model into two overarching categories—data scope and organizational data governance—which provides a clearer understanding of the complexities involved. Our examination of the conceptual framework unveils profound insights into the dimensions of data scope and data governance.

In the data scope layer, three key dimensions are highlighted. The first one is the extension of the boundaries of traditional data, surpassing them in terms of coverage, resolution, and granularity. It offers a nuanced understanding of issues such as air pollution and traffic flows. However, while expanding these data boundaries is essential, there is a need to ensure data aligns with monitoring objectives. Which brings us to our second dimension, the need for the data to be fit for purpose. Challenges include increased data processing demands and the importance of understanding local government's data needs. Alignment with the SDGs is crucial, but projects often lack awareness and struggle with the complex nature of the SDG Framework. A third dimension revolves around the representativeness of the data being collected. Inclusivity is crucial to avoid biases and align with the SDG principle of "Leave No One Behind." Challenges arise in achieving diversity, with current trends involving economically privileged individuals. The integration of underrepresented groups significantly enhances the richness of citizen science data, contributing to the democratization of knowledge crucial for informed policy formulation. Collaborations with intermediaries and educational institutions are suggested to enhance representativeness.

Regarding the data governance layer, there are two main components: the data process and data management. Within the data process, our analysis highlights the challenges of sustaining citizen science data projects beyond one-off initiatives. Often financial constraints lead to the temporal nature of data collection, presenting hurdles that require careful planning and resource allocation. Although there are examples of projects that successfully transitioned into sustainable service providers for local governments. The importance of citizen engagement is underscored, acknowledging the challenge of managing engagement fatigue. Furthermore, ethical and legal considerations are highlighted as crucial, emphasizing transparency, informed consent and clear guidelines for data ownership in citizen science data projects.

In the realm of data management, adherence to FAIR principles is identified as key to facilitating collaboration and ensuring the quality and usability of generated data. However, challenges persist in standardizing data practices, publication in open formats, achieving interoperability, and data integration with other existing data sources. Regarding data quality, our experts emphasize that not only the actual quality of citizen science data is a recurring concern, but they underscore the need for expertise in data interpretation, with higher governments in a supporting role.

Based on qualitative research, we enhanced the conceptual model for the BCR with data scope characteristics, including data extending traditional boundaries, the fitness for a specific objective and ensuring the representativeness of the collected data. The necessary design requirements are situated within the domain of data governance, including considerations related to the sustainability of data collection processes, ethical and legal dimensions, standardized data publication, data accessibility, data integration and data quality assurance.

Our results provide local stakeholders and in particular policymakers practical insights into how the barriers for uptake of citizen-generated data into local SDG monitoring of environmental indicators can be overcome. As such, we want to improve the uptake of citizen-generated data projects into Environmental SDG monitoring in the BCR.

The findings underscore the multifaceted nature of integrating citizen science data into policy, highlighting the need for a holistic and adaptive approach. Here, "holistic" refers to the inclusive consideration of diverse dimensions within the framework, while "adaptive" signifies the recognition that contextual factors play a crucial role, thereby acknowledging the limited generalizability. Effectively integrating citizen science data into SDG monitoring at the local level necessitates addressing challenges related to data scope, data governance and context-specific considerations.

In future research endeavors, this framework can be employed to substantiate the practical impact and iteratively refine the underlying framework. Additionally, its applicability may extend to other geographical territories or be adapted for application to citizen-generated data sources directed towards non environmental SDG indicators.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

KB: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. LV: Methodology, Supervision, Writing–review and editing. CV: Supervision, Writing–review and editing. LT: Supervision, Writing–review and editing. RH: Supervision, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research was done in the context of the PhD project titled "SDG in ACTION—Data Driven Coaching of Local Communities in Achieving SDG Goals" which is subsidized by the Brussels Capital Region—Innoviris.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abraham, R., Schneider, J., and Vom Brocke, J. (2019). Data governance: a conceptual framework, structured review, and research agenda. *Int. J. Inf. Manag.* 49, 424–438. doi:10.1016/j.ijinfomgt.2019.07.008

Alarcon Ferrari, C., Jönsson, M., Gebreyohannis Gebrehiwot, S., Chiwona-Karltun, L., Mark-Herbert, C., Manuschevich, D., et al. (2021). Citizen science as democratic innovation that renews environmental monitoring and assessment for the sustainable development goals in rural areas. *Sustainability* 13 (5), 2762. doi:10.3390/su13052762

Ballerini, L., and Bergh, S. I. (2021). Using citizen science data to monitor the Sustainable Development Goals: a bottom-up analysis. *Sustain. Sci.* 16 (6), 1945–1962. doi:10.1007/s11625-021-01001-1

Bowser, A., Cooper, C., De Sherbinin, A., Wiggins, A., Brenton, P., Chuang, T.-R., et al. (2020). Still in need of norms: the state of the data in citizen science. *Citiz. Sci. Theory Pract.* 5 (1), 18. doi:10.5334/cstp.303

Bryman, A. (2012). *Social research methods*. Fourth edition. Oxford University Press.

Ciambra, A., Fernández, A., Calvete, A., Miraglia, M., and Herrera-Favela, L. (2020). *Guidelines for voluntary local reviews: Volume one—a comparative analysis of existing VLRs* (volume 1; guidelines for voluntary local reviews). United cities & local governments (UCLG) and UN-habitat. https://unhabitat.org/sites/default/files/2020/07/uclg_vlrlab_guidelines_july_2020_final.pdf.

Cities Alliance (2015). *Sustainable Development Goals and Habitat III: Opportunities for a Successful New Urban Agenda* [Discussion Paper]. Cities Alliance. Available at: https://www.citiesalliance.org/sites/default/files/Opportunities%20for%20the%20New%20Urban%20Agenda.pdf.

De Rijck, K., Schade, S., Rubio, J.-M., and Van Meerloo, M. (2020). Best practices in citizen science for environmental monitoring: commission staff working document. [75pp.]. doi:10.25607/OBP-1779

EC JRC (2020). *Scaling up citizen science: what are the factors associated with increased reach and how to lever them to achieve impact*. Luxembourg: Publications Office. doi:10.2760/00926

Fraisl, D., Campbell, J., See, L., Wehn, U., Wardlaw, J., Gold, M., et al. (2020). Mapping citizen science contributions to the UN sustainable development goals. *Sustain. Sci.* 15 (6), 1735–1751. doi:10.1007/s11625-020-00833-7

Fritz, S., See, L., Carlson, T., Haklay, M., Oliver, J. L., Fraisl, D., et al. (2019). Citizen science and the united Nations sustainable development goals. *Nat. Sustain.* 2 (10), 922–930. doi:10.1038/s41893-019-0390-3

Fritzenkötter, J., Hohoff, L., Pierri, P., Verhulst, S., Young, A., and Zacharzewski, A. (2022). Governing the environment-related data space. *SSRN Electron. J.* doi:10.2139/ssrn.4250166

Greene, S., and Meixell, B. (2017). Hacking the sustainable development goals. Can US cities measure up? *[Research report]*. Washington, D.C.: Urban Insitute. Available at: https://www.urban.org/sites/default/files/publication/92946/2001490-hacking-the-sustainable-development-goals-corrected.pdf.

Hecker, S., Wicke, N., Haklay, M., and Bonn, A. (2019). How does policy conceptualise citizen science? A qualitative content analysis of international policy documents. *Citiz. Sci. Theory Pract.* 4 (1), 32. doi:10.5334/cstp.230

Herzog, C., Handke, C., and Hitters, E. (2019). "Analyzing talk and text II: thematic analysis," in *The palgrave handbook of methods for media policy research*. Editors H. Van Den Bulck, M. Puppis, K. Donders, and L. Van Audenhove (Springer International Publishing), 385–401. doi:10.1007/978-3-030-16065-4_22

Higgins, K., and Cornforth, J. (2015). Civil society and SDG monitoring. Harnessing civil society and citizen-generated data. Datashift. Available at: http://civicus.org/images/SDG%20Monitoring%20-%20DataShift%20Background%20Note%20-%20Final.pdf.

IEAG (2014). *Mobilising the data revolution for sustainable development*. Independent Expert Advisory Group, A World that Counts. Available at: https://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf.

Imran, A., and Agrawal, R. (2017). "Data provenance," in *Encyclopedia of big data*. Editors L. A. Schintler and C. L. McNeely (Springer International Publishing), 1–4. doi:10.1007/978-3-319-32001-4_58-1

Koedel, U., Schuetze, C., Fischer, P., Bussmann, I., Sauer, P. K., Nixdorf, E., et al. (2022). Challenges in the evaluation of observational data trustworthiness from a data producers viewpoint (FAIR+). *Front. Environ. Sci.* 9, 772666. doi:10.3389/fenvs.2021.772666

Lafortune, G., Fuller, G., Schmidt-Traub, G., Zoeteman, K., Mulder, R., and Dagevos, J. (2019). *The 2019 SDG index and dashboards report for European cities*. Tilburg: Sustainable Development Solutions Network (SDSN) & Brabant Center for Sustainable Development (Telos, Tilburg University. *(prototype version)* [Prototype Version]. Available at: https://s3.amazonaws.com/sustainabledevelopment.report/2019/2019_sdg_index_euro_cities.pdf.

Lämmerhirt, D., Jameson, S., and Prasetyo, E. (2016). *Making Citizen-Generated Data Work: towards a framework strengthening collaborations between citizens, civil society organisation, and others*. Mechelen: Data Shift & Open Knowledge Institute. Available at: https://www.researchgate.net/publication/311654834_Making_Citizen-Generated_Data_Work_Towards_a_framework_strengthening_collaborations_between_citizens_civil_society_organisation_and_others#fullTextFileContent.

MacFeely, S. (2019). The big (data) bang: opportunities and challenges for compiling SDG indicators. *Glob. Policy* 10 (S1), 121–133. doi:10.1111/1758-5899.12595

Mahajan, S., Chung, M.-K., Martinez, J., Olaya, Y., Helbing, D., and Chen, L.-J. (2022). Translating citizen-generated air quality data into evidence for shaping policy. *Humanit. Soc. Sci. Commun.* 9 (1), 122. doi:10.1057/s41599-022-01135-2

McKinley, D. C., Miller-Rushing, A. J., Ballard, H. L., Bonney, R., Brown, H., Cook-Patton, S. C., et al. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol. Conserv.* 208, 15–28. doi:10.1016/j.biocon.2016.05.015

Moczek, N., Voigt-Heucke, S. L., Mortega, K. G., Fabó Cartas, C., and Knobloch, J. (2021). A self-assessment of European citizen science projects on their contribution to the UN sustainable development goals (SDGs). *Sustainability* 13 (4), 1774. doi:10.3390/su13041774

Müller, M., Lorenz, J., Voigt-Heucke, S., Heinrich, G., and Oesterheld, M. (2023). Citizen science for the sustainable development goals? The perspective of German citizen science practitioners on the relationship between citizen science and the sustainable development goals. *Citiz. Sci. Theory Pract.* 8 (1), 34. doi:10.5334/cstp.583

Nascimento, S., Iglesias, J., Owen, R., Schade, S., and Shanley, L. (2018). "Citizen science for policy formulation and implementation," in *Citizen science: innovation in open science* (London: Society and Policy). Available at: https://www.researchgate.net/publication/330976243_Citizen_science_for_policy_formulation_and_implementation_Innovation_in_Open_Science_Society_and_Policy.

OECD (2020). A territorial approach to the sustainable development goals: synthesis report. *OECD*. doi:10.1787/e86fa715-en

Oosterhof, P. D. (2018). Localizing the SDGs to Accelerate the Implementation of the 2030 Agenda for sustainable development: the current State of sustainable development goal Localization in Asia and the pacific *(0 ed., governance briefs) [governance briefs]*. Manila: Asian Development Bank. doi:10.22617/BRF189612

Phillips, T. B., Parker, A., Bowser, A., and Haklay, M. (2021). "Publicly generated data: the role of citizen science for knowledge production, action, and public engagement,". *Closing the knowledge-implementation gap in conservation science*. Editors C. C. Ferreira and C. F. C. Klütsch (Springer International Publishing), 4, 83–107. doi:10.1007/978-3-030-81085-6_4

Reddy, P. (2016). Localising the sustainable development goals (SDGs): the role of local government in context. *Afr. J. Public Aff.* Vol. 9 (Number 2). Available at: https://repository.up.ac.za/bitstream/handle/2263/58190/Reddy_Localising_2016.pdf?sequence=1&isAllowed=y.

San Llorente Capdevila, A., Kokimova, A., Sinha Ray, S., Avellán, T., Kim, J., and Kirschke, S. (2020). Success factors for citizen science projects in water quality monitoring. *Sci. Total Environ.* 728, 137843. doi:10.1016/j.scitotenv.2020.137843

Sauermann, H., Vohland, K., Antoniou, V., Balázs, B., Göbel, C., Karatzas, K., et al. (2020). Citizen science and sustainability transitions. *Res. Policy* 49 (5), 103978. doi:10.1016/j.respol.2020.103978

Schleicher, K., and Schmidt, C. (2020). Citizen science in Germany as research and sustainability education: analysis of the main forms and foci and its relation to the sustainable development goals. *Sustainability* 12 (15), 6044. doi:10.3390/su12156044

SDSN, & IEEP (2019). *The 2019 europe sustainable development report: towards a strategy for achieving the sustainable development goals in the European union*. Paris and Brussels: Sustainable Development Solutions Network and Institute for European Environmental Policy. Available at: https://ieep.eu/wp-content/uploads/2022/12/Europe-Sustainable-Development-Report-2019.pdf.

Serrano, F., Sanz, F., Pelacho, M., Turbé, A., Barba, J., Mudgal, S., et al. (2018). *Citizen science for environmental policy: development of an EU wide inventory and analysis of* *selected practices*. Luxemburg: Publications Office. Available at: https://data.europa.eu/doi/10.2779/961304.

Sharifi, A., Allam, Z., Bibri, S. E., and Khavarian-Garmsir, A. R. (2024). Smart cities and sustainable development goals (SDGs): a systematic literature review of co-benefits and trade-offs. *Cities* 146, 104659. doi:10.1016/j.cities.2023.104659

UNDG (2014). *Localizing the post-2015 development agenda: dialogues on implementation*. New York: United Nations Development Group. Available at: https://uclg.org/sites/default/files/dialogues_on_localizing_the_post-2015_development_agenda.pdf.

UNEP (2021). *Measuring progress: environment and the SDGs*. Nairobi: United Nations. doi:10.18356/9789210011211

United Nations (2016). *The sustainable development goals report 2016*. United Nations. Available at: https://unstats.un.org/sdgs/report/2016/the%20sustainable%20development%20goals%20report%202016.pdf.

United Nations (2020). *Measuring and monitoring progress towards the sustainable development goals*. United Nations.

Van Audenhove, L., and Donders, K. (2019). "Talking to people III: expert interviews and elite interviews," in *The palgrave handbook of methods for media policy research*. Editors M. Puppis, H. Van den Bulck, L. Van Audenhove, and K. Donders (Palgrave Macmillan), 179–197. Available at: https://link.springer.com/chapter/10.1007/978-3-030-16065-4_10.

van den Homberg, M., and Susha, I. (2018). Characterizing data ecosystems to support official statistics with open mapping data for reporting on sustainable development goals. *ISPRS Int. J. Geo-Information* 7 (12), 456. doi:10.3390/ijgi7120456

West, S., and Pateman, R. (2017). How could citizen science support the sustainable development goals? *[Policy brief]*. Stockholm: Stockholm Environment Institute. Available at: https://mediamanager.sei.org/documents/Publications/SEI-2017-PB-citizen-science-sdgs.pdf.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3 (1), 160018. doi:10.1038/sdata.2016.18