Please cite this as: Fritz, S., et al., 2019. Citizen science and the United Nations Sustainable Development Goals. *Nature Sustainability*, *2*(10), pp.922-930. https://doi.org/10.1038/s41893-019-0390-3

Citizen Science and the United Nations Sustainable Development Goals

Steffen Fritz[1], Linda See[1], Tyler Carlson[2], Muki Haklay[3], Jessie Oliver[4,5], Dilek Fraisl[1,6], Rosy Mondardini[7], Martin Brocklehurst[8,9], Lea Shanley[10], Sven Schade[11], Uta Wehn[12], Tommaso Abrate[13], Janet Anstee[5,14], Stephan Arnold[15], Matthew Billot[16], Jillian Campbell[16], Alison Parker[17], Margaret Gold[9], Gerid Hager[1], Shan He[18], Libby Hepburn[5], Angel Hsu[19], Deborah Long[20], Joan Masó[21], Ian McCallum[1], Maina Muniafu[22], Inian Moorthy[1], Michael Obersteiner[1], Maike Weissplug[9] and Sarah West[23]

1 Ecosystem Services and Management Program, International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, Laxenburg, Austria; {fritz, see, fraisl, hager, mccallum, moorthy, oberstei}@iiasa.ac.at

2 School of Resource and Environmental Management, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada; tcarlson@sfu.ca

3 University College London; Gower Street, London, WC1E 6BT, UK; m.haklay@ucl.ac.uk

4 School of Electrical Engineering and Computer Science, Queensland University of Technology, 2 George St, Brisbane, QLD 4000, Australia; jl.oliver@hdr.qut.edu.au

5 Australian Citizen Science Association (ACSA), Australian Museum, 1 William Street, Sydney, NSW, 2010, Australia; libby@atlasoflife.org.au

6 University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

7 Competence Center Citizen Science, University of Zurich; Hirschengraben 56 8001, Zurich, Switzerland; maria.mondardini@uzh.ch

8 Citizen Science Global Partnership (CSGP); martin.brocklehurst@me.com

9 European Citizen Science Association (ECSA), Museum für Naturkunde Berlin, Invalidenstraße 43, 10115 Berlin, Germany; mg@margaretgold.co.uk

10 Nelson Institute for Environmental Studies, University of Wisconsin-Madison, 550 North Park Street, Madison, WI 53706-1491; lashanley@gmail.com; ORCID iD 0000-0001-8449-4615
11 Joint Research Center of the European Union (JRC), Via Enrico Fermi, 2749, I-21027 Ispra, Italy; s.schade@ec.europa.eu

12 IHE Delft, Westvest 7, 2611 AX Delft, The Netherlands; u.wehn@un-ihe.org

13 World Meteorological Organization (WMO), Building 7bis, Avenue de la Paix, Case postale 2300, CH-1211, Geneva, Switzerland; tabrate@wmo.int

14 CSIRO, Australia; Janet.Anstee@csiro.au

15 Statistisches Bundesamt, Destatis, Frankfurt am Main, Germany; stephan.arnold@geo-concept.de

16 UNEP Geneva; Environment House, 11 Chemin des Anémones CH-1219 Châtelaine Geneva, Switzerland; matthew.billot@un.org; campbell7@un.org

17 Woodrow Wilson Center, One Woodrow Wilson Plaza, 1300 Pennsylvania Ave., NW, Washington, DC 20004-3027; alisonjparker@gmail.com

18 CitizenScience.Asia; i@shan.blue

19 Yale-NUS College, 16 College Avenue West, Singaopore, 138527; angel.hsu@yale.edu

20 University of Dundee, 13 Perth Road, Dundee DD1 4HT, UK; d.long@dundee.ac.uk

21 Center for Centre for Research on Ecology and Forestry Applications (CREAF), Edifici C Campus de Bellaterra, 08193, Barcelona; joan.maso@uab.cat

22 Citizen Science Africa Association; mmuniafu@usiu.ac.ke

23 Stockholm Environment Institute, University of York, YO10 5NG, UK; sarah.west@york.ac.uk

Preface

Traditional data sources are not sufficient for measuring the United Nations (UN) Sustainable Development Goals (SDGs). New and non-traditional sources of data are required. Citizen science is an emerging example of a non-traditional data source that is already making a contribution. A roadmap is presented in this paper that outlines how citizen science can be integrated into the formal SDG reporting mechanisms. To succeed will require leadership from the United Nations, innovation from National Statistical Offices and focus from the citizen science community to identify those indicators to which citizen science can make a real contribution.

Main body of the paper

Advances in technology and the proliferation of data from multiple sources are providing new opportunities for monitoring and tracking the progress of the United Nations (UN) Sustainable Development Goals (SDGs)[1]. As the latest framework for assessing and monitoring the alleviation of poverty, inequalities and environmental degradation, progress on meeting the 17 SDGs is evaluated through reporting on a hierarchy of 169 targets and 232 indicators[2] (Box 1). Here we argue that data produced through ‘citizen science’[3], which is the involvement of citizens in scientific research and/or knowledge production, can complement and ultimately improve the SDG reporting process. We will demonstrate the value of using data from citizen science for the SDGs, providing concrete examples of how such data are currently being adopted in support of existing SDG indicators and their potential for contributing to other indicators in the future. We start by examining issues related to traditional data used in SDG reporting and how the emergence of new data sources can fill data gaps. We then place citizen science in the broader context of these non-traditional data streams available for SDG reporting, and highlight the value of citizen science data for the SDGs. We conclude with a roadmap containing a set of actions for mainstreaming the use of data from citizen science into official SDG reporting at the global and national levels, with a proposal for supporting activities at the local level.

Box 1: The Sustainable Development Goal (SDG) Framework

In September 2015, the United Nations 2030 Agenda for Sustainable Development, which consists of 17 Sustainable Development Goals (SDGs), was ratified[1]. This Agenda provides a framework upon which governments can implement policies and actions towards achieving these goals by 2030. The SDGs cover many areas including, among others, poverty, food security, energy, health and well-being, inequality, gender, production and consumption, urbanization and numerous environmental issues affecting land and marine ecosystems as well as climate change. Strengthening global collaborations in implementation is the subject of goal 17.

Each SDG can be elaborated as a series of targets, which can be monitored over time using 232 indicators. The global indicator framework is developed and implemented by the Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs), consisting of Member States and regional and international agencies. Data are sourced primarily from global databases maintained by international organizations, national statistical offices and other government agencies. The development of some indicators is still ongoing (see Box 2). The national governments of Member States have ownership over the SDG process and can report their progress as voluntary national reviews to the High-Level Political Forum (HLPF) on Sustainable Development, which meets annually. Established in 2012, the

HLPF has the role of reviewing the agenda, following up with Member States, and monitoring progress at the global level.

Data issues in SDG reporting

Traditional sources of data collected by national statistical offices (NSOs), e.g., through censuses, surveys, other administrative data, and in situ monitoring, as well as other government ministries and international organizations that compile global databases, currently provide the main input to SDG reporting[4]. Although valuable and necessary, these traditional sources of data, nevertheless, fall short in several ways. First, these data are costly to obtain, with population censuses ranging from hundreds of millions to 12 billion USD[5] while sample-based methods such as household surveys cost on average between 460 to 1.7 million USD depending on the type of survey used[4]. Due to these high costs, the cycle of data collection is often infrequent, and hence these traditional data sources can become quickly outdated. In addition, the data are reported at the national level so spatial variations across a country are not often captured. Finally, questions have been raised about the accuracy, openness and coverage of some official data sets[6]. For example, concerns regarding the veracity of agricultural statistics provided by member countries to the Food and Agriculture Organization of the United Nations (UN-FAO) have been discussed in the past[7].

On the question of coverage and openness, the global Open Data Inventory has compiled statistics for each country; coverage is defined as having social, economic and environmental statistics available for a minimum number of years while openness is a combined factor regarding accessibility of the data (see SI for more details). Figure 1 shows the variation in the coverage and openness of official statistical data by continent, which demonstrates large disparities across and within the continents. Moreover, greater coverage and openness are visible in continents that have more developed or affluent countries (see Figure S1 b-f for a breakdown by country). One can also see a positive relationship between coverage and openness (Figure S1a), which suggests that countries with greater capacity for data collection are also more likely to make their data open.

The gaps in statistical data coverage, identified here, need to be filled. Increasing national statistical capacity is one approach, but this requires investment. As there have been recent reported declines in the financing of sustainable development more generally, particularly in developing countries[8], this option may not occur. Another approach would be to complement official systems for SDG reporting with new, non-traditional data sources. The latter could contribute to Tier I and II indicators, but more importantly, could provide alternative methodologies and data sources for Tier III (see Box 2).

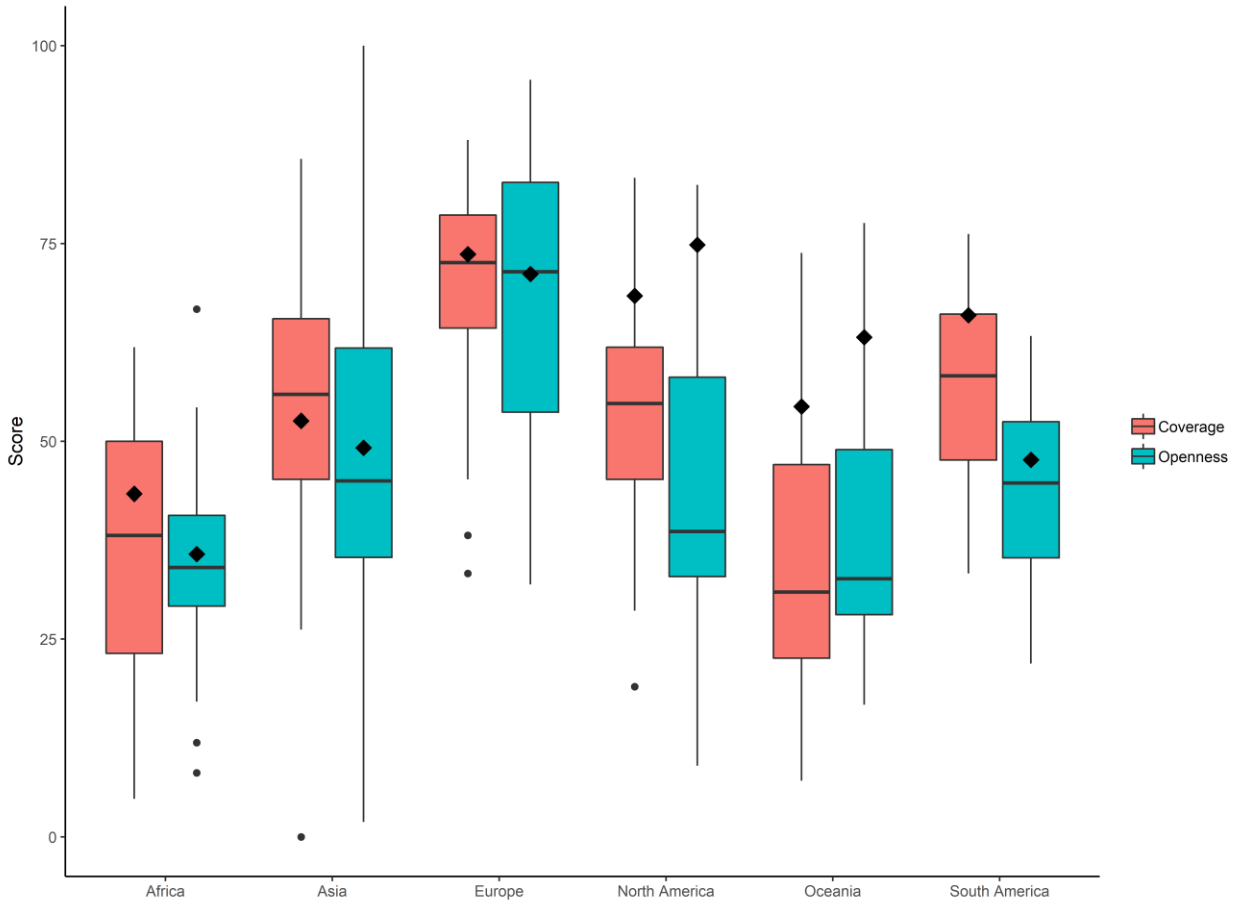

Figure 1: Boxplots of data coverage and openness by continent for 21 data categories for 2018. Grey diamonds indicated the weighted continental average by population. Data coverage refers to data availability in the last 10 years while data openness refers to five elements weighted equally, namely machine readability, use of non-proprietary formats, download options, metadata availability and terms of use. For more info, see the SI. Source: The authors, using data from the Open Data Inventory (https://odin.opendatawatch.com/data/download).

Box 2: SDG Indicator Classification by Tiers, Methodology and Current Status

Tier I: *Indicator is conceptually clear, has an internationally established methodology and standards are available. Data are regularly produced by countries for at least 50 percent of countries and of the population in every region where the indicator is relevant.*

Tier II: *Indicator is conceptually clear, has an internationally established methodology and standards are available, but data are not regularly produced by countries.*

Tier III: *No internationally established methodology or standards are yet available for the indicator, but methodology/standards are being (or will be) developed or tested.*

Source: https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/

All Tier III indicators require a work plan and methodological development to be approved by the Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs). Created by the United Nations (UN) Statistical Commission, this group is comprised of member states of the UN. The proposed methodology is then tested in one or several countries as pilots. After the methodology is refined and finalized, it can be submitted to IAEG-SDG for review and validation.

Tier I data are available for at least 50% of countries whereas Tier II data are not regularly produced. In contrast, Tier III indicators have no standardized international methodology for data collection and are either being developed or will be developed in the future. At present there are 104 Tier I, 88 Tier II, 34 Tier III indicators and 6 with multiple tiers (https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/). Countries can also propose alternative indicators based on their available data sets and specific country needs[9]. Moreover, the proportion of the Tier III indicators is much higher for the environment-related SDG indicators, with roughly 30% in the Tier III category.

New sources of data for SDG reporting

Figure 2 shows the traditional data used for SDG reporting by NSOs, other ministries and government agencies, and international organizations (top of Figure 2); these latter sources still form the main contribution to the SDGs. For example, of the 104 Tier I indicators, where the methods and data sources are established and documented in accompanying metadata, 102 indicators are based on these traditional data sources. At the same time, Figure 2 demonstrates that there are now so many other non-traditional sources of data available; these are products of the 'data revolution', which are increasingly being recognized as new and innovative sources of information for sustainable development (e.g., by the UN Secretary General's Independent Expert Advisory Group on the Data Revolution[10] and the Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs)[11]). Examples include:

- Official data sets compiled within national Spatial Data Infrastructures (such as buildings, roads and hydrological networks) and geographic information systems;
- Official sensor networks, e.g., for monitoring weather, air pollution and traffic[12];
- Commercial data sets (e.g., utility and telecommunication companies, Coca Cola's global monitoring of water quality[6] and commercial 'data philanthropy' spearheaded by UN Global Pulse - https://www.unglobalpulse.org/data-for-climate-action);
- Earth Observation (e.g., satellite imagery, LiDAR, drones); and
- Multiple sources of data generated by citizens and volunteers, both actively contributed through citizen science, or passively through social media, location-aware mobile phone data, low cost sensors and wearables.

At present, some of these non-traditional data sources are used for SDG reporting, e.g., data from air pollution monitoring stations are used by the World Health Organization (WHO) to model particulate matter (for indicator 11.6.2 *Annual mean levels of fine particulate matter (e.g. PM2.5 and PM10) in cities*), which then feeds into 3.9.1 *Mortality rate attributed to household and ambient air pollution*. The Group on Earth Observations (GEO) is spearheading efforts to support SDG monitoring with Earth Observation. In a report published in 2017, they indicate that Earth Observation can support 13 goals, 71 targets and 29 indicators[13]. However, more specifically, satellite data are currently used in Tier 1 indicators 15.2.1 *Progress towards sustainable forest management*, which uses data from the 2015 FAO-FRA (Food and Agriculture Organization-Forest Resources Assessment) exercise[14], 15.4.2 *Mountain Green Cover Index*[15] and 6.6.1 *Change in the extent of water-related ecosystems over time*, which uses the Global Surface Water Explorer data set, developed by the Joint Research Centre of the European Commission[16]. For national level reporting, see [13] for case studies in different countries.

The final source of non-traditional data shown in Figure 2 are citizen-generated data, which are defined as data produced by citizens and their organizations in monitoring issues that affect them in order to realize change[17]. The concept of citizen-generated data overlaps with many other terms[18] including citizen science, which is the focus of this paper because there are now many examples of citizen projects around the world covering a diversity of domains that can contribute to the SDGs. For example, see SciStarter.org, a search engine for citizen science projects, and a recently published inventory of citizen science activities in Europe that address environmental policies[19]. In the next section, we specifically consider data generated from citizen science and why these data have value for the SDGs.

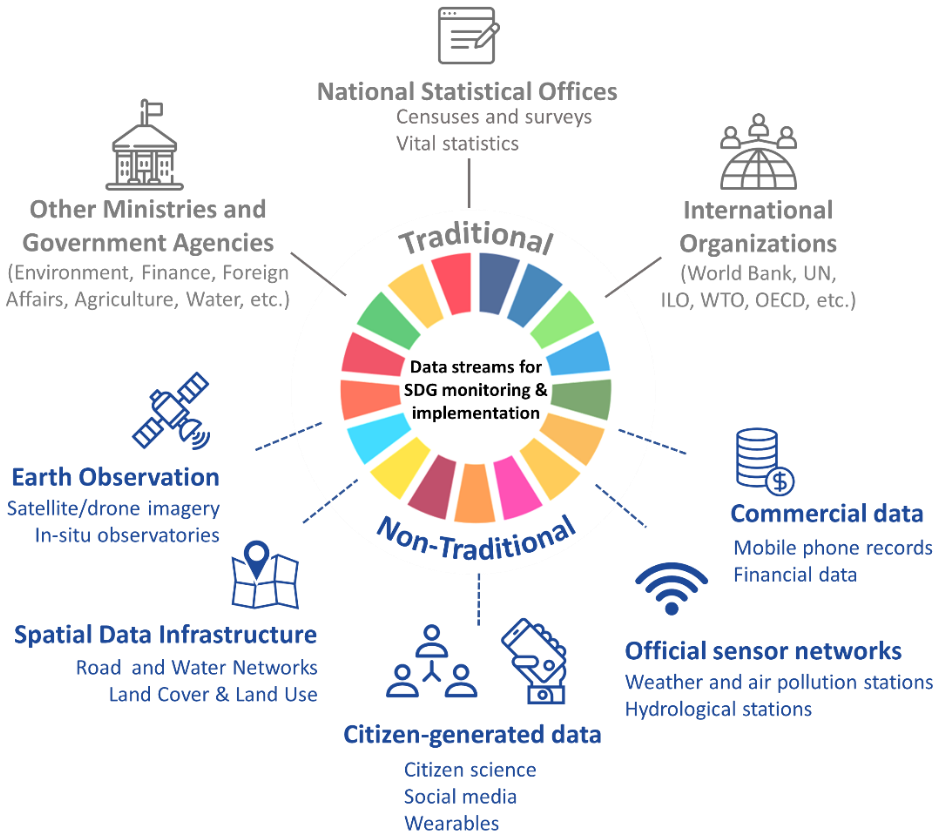

Figure 2: Traditional and non-traditional data sources available for SDG monitoring and implementation. Source: the authors.

Value of citizen science data for the SDGs

Citizen science data, like traditional data, can be characterized according to five main dimensions as shown in the first inner ring of Figure 3. Each dimension has various features, shown in the second inner ring, each of which can be shown to have value for the SDGs. The full mapping is provided in Table S1 of the SI.

The first is the *spatial dimension*, which has three characteristics: spatial reference, resolution and extent. Much of the data from citizen science projects have a spatial reference (e.g., geo-tagged photographs or location information from a mobile phone). Hence, these data can contribute to the development of spatially explicit indicators, complementing national indicator estimates. In terms of the spatial extent and resolution, citizen science projects may take place in locations that are otherwise hard to reach or more remote (e.g., through Adventure Scientists, an NGO aimed at collecting data that are difficult to obtain using a volunteer network), or they may comprise denser and more abundant observations than traditional data sources, which could lead to a greater spatial representation of an SDG indicator.

The second dimension is *temporal*, which has two main features: duration and resolution. In terms of duration, the use of regular campaigns or continuous data collection, which occurs in many citizen science projects, makes them well suited to monitoring SDG targets. At the same time, this may positively impact the temporal resolution at which the data are collected, i.e., more frequent update cycles as well as better early warning are possible with data collected by volunteers. This could then fill the temporal gap for some of the SDG indicators, which is a problem previously identified with traditional data collection.

Third is the *thematic* dimension, which is comprised of thematic subject areas, definitions and resolution. Citizen science projects cover a wide range of subject areas relevant to the SDGs (e.g., water and air quality, marine litter, biodiversity, health, gender issues). This diversity may help to address gaps in Tier II and III indicators or provide new opportunities for Tier I for various SDGs. Thematic definitions and resolution are related in that one defines the vocabulary while the second determines how detailed the vocabulary is. Some citizen science projects may use richer, more detailed, user-defined taxonomies compared to those found in more official data sets. This could potentially fill gaps in Tier II and III indicators where data sets are still missing. The challenge will be to harmonize thematic definitions and the resolution so that they are compliant with SDG indicator definitions.

Process is the fourth dimension shown in Figure 3, which has five different aspects. The first is the purpose for which the data are intended, i.e., in a way that aligns explicitly with the goals of a citizen science project or whether the data are used for another purpose, which is referred to here as implicit use. Since citizen science projects are generally not SDG-related, the data can, nevertheless, be used implicitly for SDG indicators. The second aspect is cognitive attention, which differentiates between the need for active intervention during data collection or whether the data are being collected more passively, e.g., through a low-cost sensor. Both types of data can potentially be useful as new sources of data for SDG indicators.

Data collection and processing are the next two aspects of the *process* dimension. Data collection may follow a strict sampling design or protocol that is statistically robust and could, therefore, be used to calculate an SDG indicator. Alternatively, it may be opportunistic, i.e., allowing users to collect data on the presence of a phenomenon anywhere in space and time. This type of 'presence only' data will require some data processing and modelling to infer the correct distribution of a phenomenon (e.g., of the type undertaken for determining species distributions[20]).

The fifth and last aspect is the driver of the *process*, i.e., are they contributory projects where scientists ask volunteers to collect data, or more bottom up initiatives driven by co-created and collegial processes[21]? In contributory projects, scientists can ensure that data protocols align with the requirements of SDG indicators. Co-created and collegial projects, in contrast, might promote greater ownership and contributions to key SDG indicators that are driven by the needs of the community.

The final dimension is *data management*. If data from citizen science projects are managed in line with the FAIR principles (findable, accessible, interoperable, reusable)[22], then they can be found as potential new sources for Tier II and III indicators, particularly if registered in portals such as the GEO Discovery and Access Broker; they will be openly accessible and interoperable, so that that they can be integrated with traditional data sources and across citizen science initiatives; and they will be reusable, which means they have an associated license, metadata (with provenance) and they conform to domain standards. For example, the Working Group on Data and Metadata from the Citizen Science Association is working to support interoperability between citizen science projects as well as developing and improving international data and metadata standards. Since SDG reporting is done at the national level, individual projects spread geographically can be linked to provide an integrated source of information for monitoring SDG targets. An example is the global platform eBird, which has become integrated with existing regional platforms, harmonizing data with existing projects and increasing the size and reach of the data available[23].

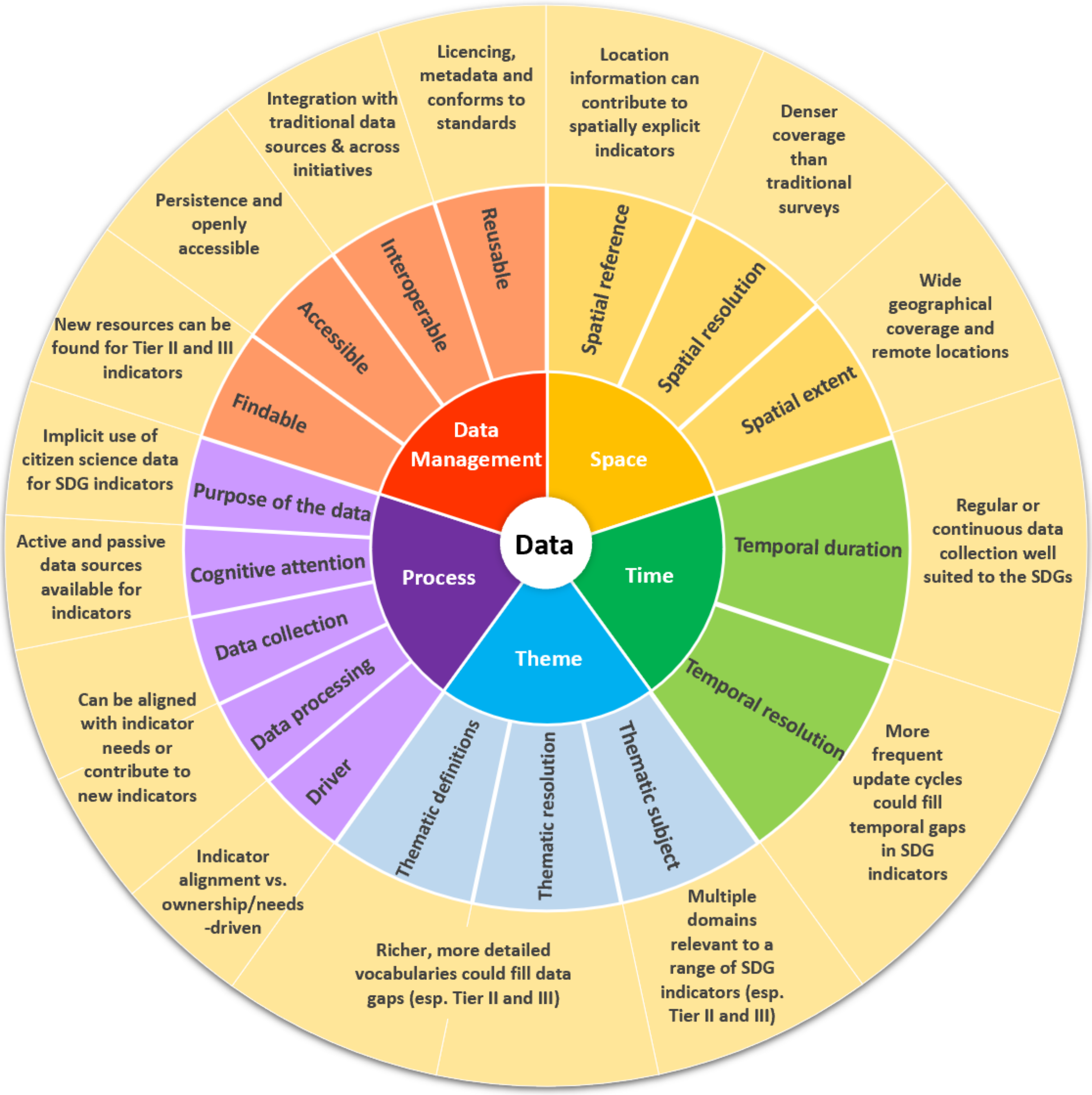

Figure 3: The five dimensions of citizen science data, their features and their value for the SDGs. Source: the authors.

Citizen science and data quality

Although we have demonstrated that citizen science data clearly have value for the SDGs, there is at least one major barrier to its use: uncertainty regarding the quality of the data. This remains one of the most discussed and researched aspects in the field of citizen science[24]. Many papers have shown, however, that citizens are able to make valuable and scientifically valid contributions that are on par with professional scientists[25]. The Mosquito Alert citizen science initiative, for example, has demonstrated that data collection is cheaper and quicker but can be obtained with the same level of accuracy as traditional methods[26].

The quality of data from citizen science can be evaluated using the same measures as any other official data (e.g., ISO19157, which is a standard used to evaluate the quality of spatial data). This includes measures such as positional and thematic accuracy, temporal currency of the data, completeness and representativeness over space and time, and whether the data are fit-for-purpose. In addition, other types of quality assurance methods specific to the nature of citizen science are needed, but many new, robust methods are now available[25,27]. Volunteer training and ongoing feedback are two of the most obvious ways to improve quality but numerous approaches have been developed, such as comparison with professionally collected data, validation by experts, peer review, filtering of outliers through automated processing, consensus-based methods including weighting by volunteer performance, and use of standardized and calibrated measurement tools[25]. Moreover, the use of AI and data mining are now increasingly being used to improve quality (e.g., by providing hints to volunteers based on automatic recognition of species from photographs)[24]. Systematic bias can be handled using the same statistical methods that are applied to data collected by professionals while approaches are being developed for handling volunteer bias[25], which is of particular relevance to citizen science.

Many citizen science projects use multiple methods to ensure high data quality. Demonstrating the quality will be a key mechanism to overcome barriers to use, as will showcasing examples of where the data are already contributing to the SDG indicators. Although not intended to be comprehensive, the examples that follow provide evidence that citizen science is already starting to be used in Tier I and II SDG reporting.

Citizen science in Tier I and II indicators

Biodiversity and conservation are two areas with strong citizen science presence and hence an area where contributions are already in evidence, e.g., Tier 1 indicator 15.5.1 *Red List Index*. Compiled by the International Union for Conservation of Nature and Natural Resources (IUCN), the Red List Index provides indicators that capture risk of extinction over time for four taxonomic groups: birds, mammals, amphibians, and corals. The organization BirdLife International compiles all the data on birds for the Red List[28], which comes from their network of scientists and more than 5,000 trained volunteers (https://www.birdlife.org/worldwide/partnership/about-birdlife). In addition, BirdLife International uses data from other relevant citizen science projects such as eBird for their list of threatened bird species and hence indicator 15.5.1; see, for example, the use of Birdata and eBird from Australia (https://birdata.birdlife.org.au/help). In the case of mammals, citizen scientists are helping identify species from photographs taken by camera traps including the crowdsourcing of threatened species[29].

Another example is the contribution of citizen science to the establishment of protected areas. More than half of the world's Key Biodiversity Areas (KBAs), which are areas designated as having international importance for biodiversity, overlap with existing protected areas[30]. KBAs contain more than 13,000 Important Bird and Biodiversity Areas (IBAs), which are established by BirdLife International using data from their volunteer network[31]. Hence citizen science contributes to additional Tier I indicators such as 15.1.2 *Proportion of important sites for terrestrial and freshwater biodiversity that are covered by protected areas, by ecosystem type*, and 15.4.1 *Coverage by protected areas of important sites for mountain biodiversity*.

At a more national level, community volunteers in the Philippines are collecting household census data on poverty, nutrition, health, education, housing, and disaster risk reduction[32], which are used by the Philippine Statistics Authority to enhance their statistics on 32 SDG indicators, including both Tier I and II[33]. Moreover, data from volunteers have been formally recognized in the Philippine Statistical Development Plan as a means for enhancing local data collection for the Philippine

Statistical System[34]. In Peru, participatory water monitoring programs enabling community involvement in data collection for watershed planning[35] have been supported by the National Water Authority of Peru, which aggregates and reports national data related to SDG 6 (clean water and sanitation). In the Andean region of Peru, local stakeholders, academic institutions and NGOs have formed the Regional Initiative for Hydrological Monitoring of Andean Ecosystems (iMHEA) to improve management of local water resources. The iMHEA network has co-developed a robust and standardized water monitoring protocol, and leverages partnerships with local universities to provide resources for training, equipment calibration, and data analysis and management[36]. Although not currently contributing, such initiatives could produce credible, supplemental data for the Tier II indicator 6.3.2 *Proportion of monitored bodies of water with good ambient water quality* while also directly supporting the achievement of SDG 6.b *Participation of local communities in improving water and sanitation management*.

Citizen science for Tier III indicators

Tier III indicators represent the greatest potential for future contributions of citizen science, both in terms of filling data gaps and in methodological developments. At present there are 34 Tier III indicators across the 17 SDGs. Here we provide examples for three of them: monitoring food waste (SDG 12); climate change (SDG 13); and marine pollution (SDG 14).

Food waste is a global issue with enormous health, economic and environmental impacts. Although tools exist to support businesses, governments and agricultural producers, citizen science approaches could be used to monitor the amount of food wasted over time, contributing to SDG target 12.3. For some European countries, robust data using standardized methods already exist[37]; these methods could be replicated in countries with data gaps while involving communities in refining these methods (e.g., determining what waste streams should be measured based on the cultural context and how best to quantify them). Technological solutions could help monitor elements of food waste such as food expiry dates and the use of ‘smart’ garbage bins.

Citizen science could also support monitoring of progress towards SDG13 through increasing human and institutional capacity to act on climate change (SDG indicator 13.3.2). An example is the climate-smart agriculture MAIS program in Brazil[38]. Small scale farmers who do not have access to expensive soil and crop monitoring services are provided with the technology and knowledge to monitor soil moisture and implement adaptive soil and land management. The proxy indicators derived from these activities include the number of farmers and growers using advice to produce locally appropriate diverse and sustainable crops despite changing conditions. In sub-Saharan Africa, farmers were equipped with low cost rain-gauges; based on the high-density data set obtained, national meteorological services were able to provide the same farmers with agrometeorological advice on the evolution of the rainy season and the most suitable farming practices[39].

Third, essential to SDG14 is the target of achieving substantial reductions in marine pollution, including nutrient pollution and marine debris in coastal waters (SDG indicator 14.1.1). Eutrophication is increasing in coastal waters, and UN Environment, the custodian agency responsible for methodological development and global reporting for this indicator, recommends the combined use of remote sensing and citizen science for large-scale monitoring with validation by citizens[40,41]. Mobile applications such as Citclops’ EyeOnWater and Earthwatch’s Freshwater Watch enable volunteers to contribute data on the color of coastal waters, which serves as a simple and accessible baseline indicator for eutrophic trends that can be used in tandem with remote sensing data[40,42]. The National Aeronautical Space Agency (NASA) in the US is exploring citizen science potential within general aviation to contribute aerial photos to assess eutrophication[43].

As with eutrophication, citizen science communities are already engaged with quantifying marine debris, contributing, for example, to the EU Plastics Strategy. At present, quantification of floating marine debris, specified in indicator 14.1.1, relies mainly on visual observations by scientists, with no standardization in approaches or an internationally agreed protocol[41]. Beyond quantifying debris that is floating, members of the public often participate in marine debris clean-ups[42,43], often recording marine debris found with the Tangaroa Blue Foundation (part of the Australian Marine Debris Initiative), OpenLitterMap, Marine Debris Tracker, and other programs. Synergies between marine and river pollution monitoring by citizen scientists could also be investigated for SDG reporting[44]. New technologies such as computer vision are being employed in combination with visual interpretation of drone imagery by citizen scientists in quantifying marine debris (e.g., in the Plastic Tide project). This could be extended to visual interpretation of imagery from larger scale aerial surveys, where volunteers could work together with NGOs and government agencies in quantifying marine debris.

Citizen science for new goals and targets

In addition to supporting the existing system of SDGs, citizen science provides opportunities to contribute to the generation of additional goals and targets where gaps can be identified. Air quality monitoring is a good example that demonstrates this potential. Currently, two SDG indicators are linked to air quality: (1) 3.9.1 *Mortality rate attributed to household and ambient air pollution*, and (2) 11.6.2 *Annual mean levels of fine particulate matter in cities*. These indicators, however, provide neither the actionable information that cities and communities need to manage their local conditions, nor do they contribute to an increased understanding of the health impacts of air pollution.

Citizen science can fill this gap through the novel application of traditional sensors such as Palmes' Diffusion Tubes[45] and the ongoing efforts to develop reliable low-cost electrochemical sensors[46]. CurieuzeNeuzen (Curious Noses) is a citizen science project on monitoring air quality in Antwerp, Belgium, using diffusion tubes. Engaging 2000 participants, the project resulted in both positive behavioral change in the participants while simultaneously driving political debate on air pollution and mobility measures[47]. Due to its success, Curious Noses has now been upscaled to the broader Flanders region of Belgium with the involvement of 20,000 participants[48]. Propeller Health is another innovative example that integrates data from sensors on Asthma inhalers with pollution information.

The current level of investment and research indicates that some air quality indicators, such as CO_2, NO_2 or Particulate Matter, may see increased adoption and use of low-cost sensors in the coming decade, especially within action-oriented monitoring schemes that can tolerate indicative levels rather than high-level accuracy for compliance purposes. Importantly, environmental protection agencies are showing commitment to the use of low-cost sensors and citizen science in air quality management at the national (e.g., the US, Netherlands and Ireland), regional (e.g., the EU), and UN level. Therefore, the opportunity exists to build a global network of projects that could be linked to a potential, new indicator, which could be in place for future global environmental monitoring efforts.

These few examples have demonstrated the value of citizen science for the SDGs, but we need to move forward, and work towards mainstreaming citizen science as an accepted methodology and source of data for SDG monitoring and reporting. Below is our suggested roadmap for building support and creating operational workflows within the UN and its member countries.

A roadmap for citizen science and the SDGs

The roadmap is organized into activities that take place at three levels, i.e., global, national and local, with interactions between these levels to reach a set of overarching goals (Figure 4). At the global level, the goal is to integrate citizen science into the formal SDG reporting process. As outlined in Box I, the SDG goals, targets and indicators have been developed by the IAEG-SDGs, in consultation with experts from the UN, civil society, businesses, academia and NSOs. This development process is still ongoing, whereby Tier III indicators are moved to Tier II and eventually Tier I, with annual reporting to the High-Level Political Forum for Sustainable Development. Each indicator is the responsibility of one or more custodian agencies, which work with experts to develop indicator methodologies. Pilot projects are then run in designated countries to demonstrate the methodology and the data collection.

To be part of this global reporting process, Tier II and III SDG indicators must be identified to which citizen science could contribute. Here we take the example of floating marine debris. As outlined above, we have already identified many initiatives that are involved in marine debris identification and clean up, providing both temporal and spatial scalability, a high level of citizen participation and project longevity. The second step is to get the custodian agencies for the indicators identified on board. In the case of marine debris, the custodian agency is UN Environment, who are currently investigating the use of citizen science for this indicator. Workshops then need to be held whereby representatives from various citizen science projects are brought together with UN Environment and experts in marine debris science, to agree a consistent set of protocols for measuring and collecting the data for SDG reporting. Clear guidance and usable tools need to be offered to citizen science projects to make their data available and fit-for-use within the SDG framework. Countries with strong national citizen science initiatives could then act as pilots for this indicator (e.g., Australia, Chile, Ireland and the UK). If successful, these efforts could be scaled up to other countries. In this way, data from citizen science could be formally brought into the SDG reporting process at the global level.

The second level of the roadmap concerns activities geared at the national level (Figure 4), particularly as responsibility for the SDG reporting process lies with national governments[49]. Here the overall goal is to build an environment of trust for the use of citizen science data by national agencies. Organizations such as the Citizen Science Global Partnership (CSGP), citizen science associations and their working groups, and current Communities of Practice (CoPs) in citizen science should work actively with NSOs to bring citizen science into the scope of official statistics. A number of activities can take place in parallel, some of which have already been initiated. For example, the CoP on citizen science and SDGs, running within the EU-funded WeObserve project (https://www.weobserve.eu/), is currently mapping existing contributions of citizen science to SDG indicators and is identifying those indicators where projects could potentially contribute in the future. This mapping can help inform activities taking place at the global level.

Another activity is to compile an inventory of examples of good practice in the use of citizen science at the national level. The acceptance of data collected by citizens in the Philippines as outlined above is a prime example[34] while Australia has developed its own "Method for Australia's SDG baseline assessment" that allows for the reporting and monitoring of quantities adapted to their specific country context[50]. For many of these "proxy" indicators, citizen participation provides the main data collection mechanism. Aligning citizen science with the priorities of decision makers at the national level in this way will increase the likelihood of its adoption in SDG monitoring and reporting in other countries. Based on this inventory, we should seek to replicate models of good practice already taking place nationally.

A third activity is to build on existing policy frameworks that advocate citizen science for decision making (e.g., the 2015 US White House Memorandum on Citizen Science; the US Crowdsourcing and Citizen Science Act, which was incorporated into Section 402 of the American Innovation and Competitiveness Act; the European Open Science Policy Agenda; the recommendations of the European Open Science Policy Platform (OSPP); action 8 of the EU roadmap to streamline environmental monitoring[51]; and the Ministerial declaration at the UN Environment Assembly at its fourth session[52], which calls for the wider use of data from citizen science). Here the citizen science community should work with decision makers in other countries around the world to craft policies that authorize, encourage, and provide guidelines for the appropriate use of citizen science for SDG monitoring.

A fourth key activity in the roadmap will be to promote dialogue on data quality, which will be the number one concern for NSOs, as well as data management including standards, metadata and interoperability. The citizen science community should work with statistical agencies and the scientific community to develop an agreed set of accepted protocols and minimum data quality standards required from citizen science projects, drawing on a large body of research that already addresses quality issues and involving relevant working groups such as the CSA Working Group on Data and Metadata. Moreover, opportunities should be created for peer-to-peer networking among statisticians to share lessons learned and best practices in using citizen science for SDG monitoring, including harmonizing data from citizen science with more traditional sources of data. These activities need to be paralleled by efforts to raise awareness of the current data quality assurance mechanisms so that perceptions (and not just procedures) change that may otherwise stand in the way of the using data from citizen science for the SDGs. At the same time, these best practices should feed into the activities taking place at the global and local levels.

In addition to data quality, interoperability is paramount to interpreting, sharing and integrating the data across different citizen science initiatives, and with traditional and non-traditional data sources (e.g., satellite imagery). The citizen science community, in collaboration with SciStater.com and other partner organizations, have been developing a new data and metadata standard for Public Participation in Scientific Research (PPSR), adapted from the Darwin Core standard for biodiversity[53]. Aligning PPSR core standards with the SDG indicators can reveal where citizen science has the potential to contribute the most, maximizing the efficiency of data infrastructure, storage and curation. Future iterations of the PPSR core standard can incorporate economic, environmental and societal dimensions of the SDGs. Cloud computing, Artificial Intelligence and other frontier technologies are also making this integration possible. The ultimate activity at the national level will be to work with NSOs to integrate data from citizen science in their statistical reporting.

The third and final level of the roadmap concerns activities that affect citizen science projects operating at the local level. The citizen science community must actively support citizen science projects through guidance and tools that will help them adhere to the FAIR data principles, while promoting the sharing and reuse of the data. Best practices developed at the national level should feed down to the local level. At the same time, the citizen science community must identify potential privacy risks and help citizen science projects adhere to applicable data privacy laws. The US White House Office of Science and Technology Policy memorandum on citizen science, issued on September 30, 2015, provides guiding principles, including providing volunteers with appropriate access to their data, ensuring meaningful engagement of the public in scientific research, and giving appropriate attribution for volunteer contributions. In addition, the European Citizen Science Association has developed a set of 10 principles for citizen science that should be followed[54]. Finally, the citizen science community should raise awareness of the SDGs among citizen science projects, encouraging them to align their goals to SDG monitoring where relevant.

Figure 4: Roadmap of activities for integrating citizen science into SDG reporting

Conclusion

SDG indicators are largely fed by traditional data from NSOs, other government ministries/official agencies, and international organizations. Yet we are now in the midst of a data revolution, with the emergence of new sources of non-traditional data that can fill the increasing demand for high-resolution spatial and temporal data. Not only can these data sources feed our models and aid in timely decision making, they should become inputs to the SDGs. Citizen science, in particular, should contribute in this increasingly data-hungry world. Although we have demonstrated that data from citizen science are already starting to contribute to some of the indicators (Tier I and II), the real potential exists in contributing to Tier III indicators, where the data and/or methodologies are still being developed. Now is the time to identify those indicators to which citizen science can make real contributions and mobilize the citizen science community to become an active part of the SDG reporting process at the global level. At the same time, much work is still needed to create a trusted environment in which citizen science data are accepted as a credible source of inputs for statistical reporting at the national level. Providing support to citizen science projects at the local level will ensure that they can contribute to SDG reporting while tapping into an opportunity for social innovation whereby citizens can help to both monitor but more importantly implement the SDGs.

Acknowledgements

Partial funding was provided by the EU's Framework 7 European Research Council project CrowdLand (No. 617754) and the EU Horizon 2020 project WeObserve (No. 776740).

Correspondence and Request for Materials

The corresponding author of the paper is Steffen Fritz, International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, A-2361 Laxenburg, Austria. Email: fritz@iiasa.ac.at.

References

1. *UN General Assembly Transforming our World: the 2030 Agenda for Sustainable Development* A/RES/70/1 (United Nations, 2015).
2. *Revised List of global Sustainable Development Goal Indicators* (IAEG-SDGs 2017).
3. Shanley, L. A., Hulbert, J. & Auerbach, J. *Definitions of citizen science*. https://github.com/lshanley/CitSciDefinitions (2019).
4. Espey, J. *et al. Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development* (UNSDSN 2015).
5. *Main Results of the UNECE-UNSD Survey on the 2010 Round of Population and Housing Censuses* (UNECE and UNSD 2009).
6. Hsu, A., Malik, O., Johnson, L. & Esty, D. C. Development: Mobilize citizens to track sustainability. *Nature* **508**, 33–35 (2014).
7. World Bank & FAO. *Global Strategy to Improve Agricultural and Rural Statistics* (The World Bank, 2011).
8. *Global Outlook on Financing for Sustainable Development 2019: Time to Face the Challenge* (OECD 2018).
9. Bizikova, L. & Pinter, L. *Indicator Preferences in National Reporting of Progress Toward the Sustainable Development Goals* (IISD 2017).
10. *A World that Counts. Mobilising the Data Revolution for Sustainable Development* (IAEG Secretariat 2014).
11. Arnold, S., Chen, J. & Eggers, O. *Global and Complementary (Non-Authoritative) Geospatial Data for SDGs: Role and Utilization* (Working Group on Geospatial Information, IAEG-SDGs, 2019).
12. See, L. *et al.* Using remote sensing and geospatial information for sustainable development. in *From summits to solutions: innovations in implementing the Sustainable.Development Goals* (eds. Desai, R. M., Katō, H., Kharas, H. J. & McArthur, J. W.) 172–198 (Brookings Institution, 2018).
13. *Earth Observations 2030 Agenda for Sustainable Development, V1.1. Japan Aerospace Exploration Agency (JAXA) on behalf of GEO under the EO4SDG Initiative* (GEO 2017).
14. MacDicken, K. G. Global Forest Resources Assessment 2015: What, why and how? *Forest Ecology and Management* **352**, 3–8 (2015).
15. *Metadata for indicator 15.4.2* (United Nations 2017).

16. *Monitoring our Blue Planet: First SDG indicator platform launched by Google, the JRC and UN Environment* (EC 2019).
17. *Using Citizen-Generated Data to Monitor the SDGs. A Tool for the GPSDD Data Revolution Roadmaps Toolkit* (DataShift, 2017).
18. Eitzel, M. V. *et al.* Citizen science terminology matters: Exploring key terms. *Citizen Science: Theory and Practice* **2**, 1 (2017).
19. *Citizen Science for Environmental Policy: Development of an EU-wide Inventory and Analysis of Selected Practices*. Final report for the European Commission, DG Environment under the contract 070203/2017/768879/ETU/ENV.A.3, in collaboration with Fundacion Ibercivis and The National History Musuem (Bio Innovation Service 2018).
20. Fithian, W., Elith, J., Hastie, T. & Keith, D. A. Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods Ecol Evol* **6**, 424–438 (2015).
21. Shirk, J. L. *et al.* Public participation in scientific research: A framework for deliberate design. *Ecology and Society* **17**, (2012).
22. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018 (2016).
23. Sullivan, B. L. *et al.* The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation* **169**, 31–40 (2014).
24. Lukyanenko, R., Wiggins, A. & Rosser, H. K. Citizen science: An information quality research frontier. *Inf Syst Front* (2019).
25. Kosmala, M., Wiggins, A., Swanson, A. & Simmons, B. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment* **14**, 551–560 (2016).
26. Palmer, J. R. B. *et al.* Citizen science provides a reliable and scalable tool to track disease-carrying mosquitoes. *Nature Communications* **8**, 916 (2017).
27. Wiggins, A., Newman, G., Stevenson, R. D. & Crowston, K. Mechanisms for data quality and validation in citizen science. in *2011 IEEE Seventh International Conference on e-Science Workshops* 14–19 (IEEE, 2011).
28. Burfield, I. J., Butchart, S. H. M. & Collar, N. J. BirdLife, conservation and taxonomy. *Bird Conservation International* **27**, 1–5 (2017).
29. Schuttler, S. G. *et al.* Citizen science in schools: Students collect valuable mammal data for science, conservation, and community engagement. *BioScience* **69**, 69–79 (2019).
30. Eken, G. *et al.* Key Biodiversity Areas as site conservation targets. *BioScience* **54**, 1110–1118 (2004).
31. Donald, P. F. *et al.* Important Bird and Biodiversity Areas (IBAs): the development and characteristics of a global inventory of key sites for biodiversity. *Bird Conservation International* 1–22 (2018).
32. *CBMS Accelerated Poverty Profiling (APP): Scan, Portal, and Field Operations Manual. Partnership for Economic Policy* (PEP 2016).
33. *Localizing the Sustainable Development Goals using CBMS and LOGOD. Regional Knowledge Exchange* (United Nations Development Program 2018).
34. Philippine Statistical Development Program: 2011-2017 Strategic Plan. National Statistical Coordination Board (NSCB 2011).
35. *El Monitoreo Participativo del Agua* (Autoridad Nacional del Agua, Gobierno del Perú, 2018).
36. Ochoa-Tocachi, B. F. *et al.* Impacts of land use on the hydrological response of tropical Andean catchments. *Hydrological Processes* **30**, 4074–4089 (2016).
37. *Toidujäätmete ja toidukao teke Eesti kodumajapidamistes ja toitlustusasutustes*, Stockholm Environment Institute Project Report (SEI 2015).
38. *MAIS Program: Climate-Smart Agriculture Brazil* (UNFCCC 2018).
39. Tarchiani, V. *et al.* Smallholder farmers facing climate change in West Africa: Decision-making between innovation and tradition. *Journal of Innovation Economics* **24**, 151 (2017).

40. Busch, J. A., Price, I., Jeansou, E., Zielinski, O. & van der Woerd, H. J. Citizens and satellites: Assessment of phytoplankton dynamics in a NW Mediterranean aquaculture zone. *International Journal of Applied Earth Observation and Geoinformation* **47**, 40–49 (2016).
41. *A Snapshot of the World's Water Quality: Towards a Global Assessment* (UNEP 2016).
42. Cochero, J. AppEAR: Una aplicación móvil de ciencia ciudadana para mapear la calidad de los hábitats acuáticos continentales. *Ecología Austral* **28**, 467–479 (2018).
43. Ansari, R. R. & Schubert, T. *General Aviation Citizen Science Study to Help Tackle Remote Sensing of Harmful Algal Blooms (HABs)*. (Glenn Research Center, 2018).
44. Hadj-Hammou, J., Loiselle, S., Ophof, D. & Thornhill, I. Getting the full picture: Assessing the complementarity of citizen science and agency monitoring data. *PLOS ONE* **12**, e0188507 (2017).
45. Haklay, M. & Eleta, I. On the front line of community-led air quality monitoring. in *Integrating Human Health into Urban and Transport Planning: A Framework* (eds. Nieuwenhuijsen, M. & Khreis, H.) 563–580 (Springer, 2019).
46. Clements, A. L. *et al.* Low-cost air quality monitoring tools: From research to practice (a workshop summary). *Sensors* **17**, 2478 (2017).
47. Van Brussel, S. & Huyse, H. Citizen science on speed? Realising the triple objective of scientific rigour, policy influence and deep citizen engagement in a large-scale citizen science project on ambient air quality in Antwerp. *Journal of Environmental Planning and Management* 1–18 (2018).
48. *Curious Noses Measure Air Quality across Flanders* (FLANDERS TODAY 2018).
49. *Guidelines to Support Country Reporting on The Sustainable Development Goals* (UN Development Group 2017).
50. *Tracking Australia's Progress on the Sustainable Development Goals 2018* (Australian Government 2018).
51. *Report from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Actions to Streamline Environmental Reporting*. COM(2017) 312 final (EC 2017).
52. *UNEP/EA.4/HLS.1 Ministerial declaration of the United Nations Environment Assembly at its fourth session. Innovative solutions for environmental challenges and sustainable consumption and production* (UNEP 2019).
53. Bowser, A. *et al. Citizen Science Association Data & Metadata Working Group: Report from CSA 2017 and Future Outlook*. (Wilson Center, 2017).
54. Robinson, L., Cawthray, J., West, S., Bonn, A. & Ansine, J. Ten principles of citizen science. in *Citizen science: innovation in open science, society and policy* (eds. Hecker, S. et al.) 27–40 (UCL Press, 2018).