



ACTING LOCALLY MONITORING GLOBALLY?

**HOW TO LINK
CITIZEN-GENERATED DATA
TO SDG MONITORING**

Shazade Jameson
Danny Lämmerhirt
Eko Prasetyo



Shazade Jameson is an independent social science researcher on governance issues.

Danny Lämmerhirt is research coordinator and researcher at Open Knowledge International.

Eko Prasetyo is a practitioner focusing on ICT4D project implementation and evaluation.





**ACTING
LOCALLY
MONITORING
GLOBALLY?**

—
HOW TO LINK
CITIZEN-GENERATED DATA
TO SDG MONITORING



ACKNOWLEDGEMENTS

We would like to recognise our gratitude to the following people whom we interviewed, or otherwise drew inspiration from in this project.

Pieter Franken, Safecast	Helen Turvey, Shuttleworth Foundation
Azby Brown, Safecast	David Kennewell, Hydrata
Paul Vicks, Patients Like Me	Kingsley Purdam, Manchester University
Nicolas Kayser-Bril, Journalism++	Yantisa, Akhadi, Humanitarian OpenStreetMaps Team Indonesia
Dr. Claudia Abreu Lopes, Africa's Voices Foundation	Jennifer Walker, Code for South Africa
Rainbow Wilcox, Africa's Voices Foundation	Dimitri Katz, DevelopmentCheck
Mario Roset, Wingu	Carolin Gerlitz, Siegen University
Jennifer Bramley, mySociety	Bhaskar Mishra, UNICEF Tanzania
Jennifer Gabrys, Goldsmiths University	Karen Rono, Development Initiative
Dr. Saliou Niassy, Land Matrix Initiative	Daniel Lombraña Gonzales, SciFabric
Zoe Fairlamb, WeFarm	Claire Rhodes, Cafédirect Producers Foundation
Christine Richter, University of Twente	Mojca Cargo, GSMA
Linnet Taylor, Tilburg Institute of Law and Technology	



TABLE OF CONTENTS

FOREWORD	7
EXECUTIVE SUMMARY	8
RECOMMENDATIONS	9
THE ROLE OF RELEVANCE AND THE POLITICS OF DATA	10
INTRODUCTION	11
1 MAPPING CGD TO THE SDG INDICATORS	14
CGD projects for local action	
Often use different framings than the SDGs	15
How CGD projects currently link to the SDGs	16
CGD mostly address SDG targets, not indicators	17
CGD needs to be linked with other data, usually at a national level	17
Government frameworks and strategies are difficult to monitor at SDG level	19
Data for the SDG indicators is often tangential to CGD project aims	19
Not all CGD addresses developing countries	20
2 COMPARING AND LINKING CGD	21
3 AGGREGATING CGD FOR THE SDGS	23
Aggregating CGD flood risk data for the SDG indicators	23
RECOMMENDATIONS	26





FOREWORD

Citizens and their organisations create data as a direct reflection of their issues. This citizen-generated data (CGD) yields the potential to tell counter-narratives of sustainable development, and truly make all voices count. Yet, critics argue that citizen-generated data lacks rigorousness and is not representative; as a result, the recurring question is what are the factors that can help or hinder the usability of CGD, increase its uptake and help drive the action that it aims for.

There are different paradigms of using data for monitoring or using data for action, different perceptions around similar topics, and different data quality requirements at different scale levels. For instance, monitoring is only one aspect in the larger cycle of human decision-making including agenda setting, designing, and implementing solutions. These actions are equally important to drive sustainable development as monitoring, but these issues get little attention in current debates. Therefore, it is important to understand which forms of action can be informed by citizen-generated data, and when and how monitoring of the higher level indicators such as the SDGs can be useful.

In order to properly zoom in on and untangle these differences, we present two research reports that work together in tandem. This piece, 'Acting Locally, Monitoring Globally?', explains how citizen-generated data can help monitoring the SDGs, discussing the challenges and opportunities that arise as the data moves from being used for action to being used for monitoring at a higher-scale level. The tandem piece, 'From Evidence to Action', focuses on factors that help make CGD relevant and actionable. It emphasises that in order for data to inform decisions around sustainable development, the data must be catered to different stakeholders in different forms with different aspects of data quality.

The research series was commissioned by DataShift, an initiative that builds the capacity and confidence of civil society organisations to produce and use data, especially citizen-generated data, to drive sustainable development. It also builds on former research by Open Knowledge International on what can be done to make the "data revolution" more responsive to the interests and concerns of civil society.¹

¹ Gray, J. (2015). Democratising the Data Revolution: A Discussion Paper. Open Knowledge. Available at: <http://blog.okfn.org/2015/07/09/democratising-the-data-revolution/>, Gray, J., Lämmerhirt, D. (2015): Changing What Counts: How Can Citizen-Generated and Civil Society Data Be Used as an Advocacy Tool to Change Official Data Collection. Available at: <http://civicus.org/thedatashift/wp-content/uploads/2016/03/changing-what-counts-2.pdf> as well as Gray, J. and Lämmerhirt, D. (forthcoming): Data And The City. How Can Public Data Infrastructures Change Lives in Urban Regions?

EXECUTIVE SUMMARY

The Sustainable Development Goals (SDGs) are an effort to “end all forms of poverty, fight inequalities and tackle climate change” by 2030, worldwide. Much attention has been paid to the question how to measure and monitor progress around the goals. Monitoring is one out of several elements of decision-making, serving to evaluate progress around a target over time. As such the SDGs are an important benchmark how sustainability in nation states develops. A rallying call is that SDG indicators must not only collect national aggregates, but also recognise marginalized and underrepresented communities in order to “leave no one behind” and “make all voices count”.

Citizen-generated data (in short: CGD) is data actively created by citizens and their organisations. It is produced to monitor, demand or drive change around issues that are important to them, often collected on the ground and in local contexts. As such, CGD yields the potential to foreground the issues of disadvantaged communities in underreported regions. Yet, if the SDGs aim to leave no one behind, and if CGD can provide contextual information about marginalized groups, the question is whether and how CGD can be used for the SDGs.

Currently there is often a mismatch between citizen-generated data (CGD) & SDG monitoring:

- CGD has **more focus on local action** than high-level policy-making
- CGD tends to **focus on SDG targets** rather than indicators, using different units of analysis
- CGD can deliver contextual information to drive progress around sustainable development, but the **causal relationship between CGD and how it connects to the SDG indicators is often unclear**
- It is often **the action that is undertaken as a result of CGD** that will contribute the most to drive progress around sustainable development. Yet this local action needs to be scaled up to tie-in to national SDGs

RECOMMENDATIONS

In order to use CGD for SDG monitoring, we recommend that civil society and governments:

- **Link and aggregate multiple data** sources at different spatial scales, depending on the governance level addressing the issue
- Encourage **quantitative monitoring data** to be part of CGD projects by creating incentives to collect monitoring data even if only a secondary project aim
- Build methodological capacities around **standardisation** of both data types and collection methods to facilitate comparison and foster trust in the data
- Work with and build on **existing government processes** to help align interests
- Use CGD as **longitudinal data snapshots** to measure across time in order to contribute to trend analyses

CGD can significantly contribute to future possibilities in expanding SDG monitoring, both in terms of collaborating with governments as well as civil society engaging in shadow monitoring:

- CGD can help **combine quantitative with qualitative data** to increase the validity of the SDG indicators
- CGD can deliver **parts of the necessary data, or at least rich contextual information**, whenever indicators measure a causal relationship
- CGD can often be flexibly applied across cross-cutting SDG targets and themes, thereby overcoming silo thinking and creating **alternative indicators**
- **Use the SDGs as a common framing** to build connections across initiatives and facilitate communication and prioritisation of issues
- Because a lot of CGD focuses on building local capacity, there is a lot of room for CGD to **contribute to progress on SDG 16**

THE ROLE OF RELEVANCE AND THE POLITICS OF DATA

CGD projects are designed to solve an issue by addressing and engaging with stakeholders. Only if CGD is relevant, it can steer behaviour, or enable decision-making of those stakeholders. **The question is paramount, for whom is which type of information useful, when trying to scale locally collected data to regional and national levels.** It is important to recognize that **the governance around an issue defines the necessary level of [spatial] data aggregation required.** For instance: National government bodies may be responsible to allocate money to regions for water point constructions. Responsibility for their maintenance may reside with local districts. While national government needs comparative data across regions to allocate infrastructure investments, local districts need hyperlocal waterpoint information. An ongoing monitoring has to be conducted on a district-by-district basis, built on hyperlocal maintenance data. Before intending to scale CGD, it should be asked **for whom does data at another scale matter.**

The very process of creating CGD is born out of priorities, what to measure and how. In some cases this resolves in incompatibility of CGD across local contexts. To link CGD to the SDGs would require besides building human and technical capacity, to collect data across regions. To do so, it would be important to **ensure interoperability of CGD collected in different contexts, and driven by different interests and issues.** In order to scale CGD projects, collective data standards, metadata, or other documentation can be developed rendering CGD comparable. It should be noted that **standardisation is a highly political process** potentially evening out local differences of data. There is also a **potential conflict between local interests in data, and interests in scaling the data.** Linking CGD to the SDGs is therefore to a large extent a matter of understanding power dynamics, creating participatory processes to define what counts as evidence, and aligning interests.

INTRODUCTION

Recently, much attention has been paid to the Sustainable Development Goals (SDGs), their targets, and how they measure sustainable development through indicators. The SDGs are universal efforts to “end all forms of poverty, fight inequalities and tackle climate change” by 2030. The 17 Goals set out 169 targets for all countries, and “key to success on one will involve tackling issues more commonly associated with another”.² The SDGs are characteristic of the global policy landscape in that they are a high-level monitoring initiative with clearly framed indicators to measure progress on a series of outputs.

There are three elements that make monitoring the SDGs difficult. First, it requires monitoring and evaluation data that is either not regularly produced (Indicator Tier 2) or data that does not exist at all (Indicator Tier 3)³. Second, as an accountability mechanism, whilst the SDG indicators incentivise action for sustainable development, the bureaucratic exercise of monitoring is not necessarily the same as action towards sustainable development. This means that data driving action might not be the same as data evidencing progress on the indicators. Third, the SDG monitoring effort requires working across three scales; countries must translate the global agenda into national priorities and national methods of monitoring, and then projects for local action need to link up to the higher level monitoring effort. Whilst this multi-scalar approach seemingly makes it harder to measure progress across countries, it encourages national ownership and helps local projects identify gaps to either work with government to fill gaps on prioritised issues, or engage in shadow monitoring on important issues the government hasn’t or can’t prioritise.

As part of the movement for evidence-based sustainable development, citizen-generated data (CGD) can significantly contribute to the SDG monitoring effort by filling data gaps and giving more contextual understanding to indicators. **CGD is data actively created by citizens and their organisations to monitor, demand or drive change** around issues that are important to them. As a direct reflection of citizens’ issues, it is a strong voice driving action around inclusive and sustainable development, particularly in disadvantaged communities of the world where adequate data is lacking.

² UNDP <http://www.sdgfund.org/mdgs-sdgs>

³ See also: <http://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-03/Provisional-Proposed-Tiers-for-SDG-Indicators-24-03-16.pdf>

How can CGD best support the monitoring effort around the SDGs? This is not a simple process of addition, and it requires several intermediary steps of data processing and aggregation at different levels of governance. Indicators have different meanings depending on their target audience, and the SDG indicators are framed for national level monitoring, whereas the primary focus of CGD is usually to stimulate local action.

RESEARCH QUESTIONS

The report addresses and is structured around the following questions: to what extent are the SDGs relevant for CGD initiatives? How do current CGD initiatives link to the SDGs? What are the methodological challenges of using CGD as a data source to evidence progress around an SDG indicator? How can CGD data be aggregated to inform progress around the SDGs? What possibilities for SDG monitoring does CGD open up?

FOR WHOM IS AN INDICATOR RELEVANT?

The SDG targets and indicators stem from the global policy culture of results-based accountability, asking who is held accountable for what, using commonly understood criteria allowing auditors to evaluate performance and compliance. Much of the effort around collecting data for the SDGs is about monitoring, measuring progress over time using the indicators as a mechanism to reach the goals. Indicators measure input, processes, outputs, or outcomes, and are used in different ways, from accountability to performance and quality measurements.

Whether data is used for reporting, monitoring, or decision-making depends on the audience. The choice of how to use which indicator depends on what shall be reported to whom. In general, indicators help inform decision makers about the design and evaluation of programs and resource allocation.⁴ This means that **indicators have different meanings depending on who plays the roles of auditor and auditee.**⁵ For one it may be a measure of progress, for another a standard of quality, or for another a burden to provide evidence. For indicators to be actionable and relevant, they have to be designed so they stimulate a certain behaviour of an auditee. Thus, the choice of which data to capture to produce which indicators

4 See also: Horsch, K. (1997), Indicators: Definition and Use in a Results-Based Accountability System. Available at: <http://www.hfrp.org/publications-resources/browse-our-publications/indicators-definition-and-use-in-a-results-based-accountability-system>

5 For a good introduction into use cases, see van den Berghe, W. (1997). Indicators in Perspective. The Use of Quality Indicators in Vocational Education and Training. European Centre for the Development of Vocational Training. Available at: http://schoolinclusion.pixel-online.org/files/training_package/mod4/en/Document%2003%20Van%20Den%20Berghe.pdf

is important, because it has implications on how individuals are incentivised.⁶ The SDGs are a voluntary framing of relevant sustainability areas governments can subscribe or aspire to. While they are not an accountability or compliance mechanism, the way they frame goals, targets and indicators is likely to shape the debate about how sustainable development can be achieved.⁷

However, the indicators are but a means to an end. The main purpose is to reach the SDGs (and sustainable development more broadly), which first and foremost requires action. **Monitoring that action is only a secondary step**, whether it be to evaluate progress after a project, or to establish baselines to see how to progress further. Particularly because action often happens at a local scale in the realm of the everyday, whereas monitoring is at a higher policy level, action and global monitoring have different priorities which shape incentives at different levels of governance⁸. This means the types of questions being asked and the data needed to answer them will be different. Monitoring at scales that are higher than action also requires data to be aggregated, which can lead to methodological questions of comparability and coverage, as well as challenges of relevance and losing necessary contextual nuances.⁹

Naturally it is important to consider both local action and high-level monitoring. However, the fact that CGD is more relevant for local action is an important dimension in the context of linking CGD to the SDGs. Intrinsicly, CGD can have a valuable contribution to monitoring, though there monitoring is but one aspect of the value that CGD provides. This also depends on the goals of specific CGD projects; whilst some aim to use data as instrumental for action on an issue, such as providing feedback on government services, other aim specifically at monitoring an issue, such as tracking corruption during elections. The different primary purposes for creating data ultimately shape what databases looks like and to what extent they can be repurposed . Some of the challenges and opportunities that arise in using CGD for higher-level monitoring as explored in the following section.

6 See also: Shaw, J., Taylor, R., Dix, K. (2015). Uses and Abuses of Performance Data in Healthcare. Available at: <http://www.drfooster.com/updates/recent-publications/uses-and-abuses-of-performance-data-in-healthcare/>

7 Research demonstrated that the Millennium Development Goals had significant impact in creating tunnel visions for policy making and reducing the understanding of sustainability to very narrowly defined aspects covered in the MDGs. See for further details: Fukuda-Parr, S., Yamin, A. E., Greenstein, Y., (2014): The Power of Numbers: A Critical Review of Millennium Development Goal Targets for Human Development and Human Rights. Journal of Human Development and Capabilities. Available at: https://cdn2.sph.harvard.edu/wp-content/uploads/sites/5/2014/05/Fukuda-Parr.Yamin_Greenstein.PON_2014.pdf

8 Clearly action can happen at high policy levels as well. Because these reports focus on the potential of citizen-generated data, which is usually at the local level, we emphasise the local scale level.

9 For further details see: Lämmerhirt, D., Jameson, S. Prasetyo, E. (2017): From Evidence to Action.

MAPPING CGD TO THE SDG INDICATORS

Citizen-generated data is a means to drive progress on issues citizens and their organisations care about. In the wake of the data revolution, citizen-generated data is seen as an alternative data source to gain fine-grained, hyperlocal information. This section discusses the ways in which citizen-generated data relates to the SDGs. It shows that citizen-generated is rarely immediately usable for monitoring individual SDG indicators. Instead, it 1) provides partial data which has to be complemented by additional data, 2) offers contextual information around indicators, 3) can be applied to several targets and indicators tackling issues more holistically. These three points help answer to the critique of the SDGs as measuring progress in silos, with **overlaps, interdependencies, constraints, reinforcements, and trade-offs** across targets and indicators. With the SDG framework, it is sometimes unclear what the difference is unclear between what is an outcome goal to be achieved and what is a goal for an enabling factor to move towards the goals. The same applies to CGD in relation to the SDGs.¹⁰

The Agenda 2030 is metrified through the Sustainable Development Goals which are split into targets and indicators. Goals describe larger outcomes, targets represent the process to achieve outcomes, and indicators measure output¹¹ (see Table 1). This is visible in the way they are phrased: the targets are about doing and they start with a verb (e.g. ensure, broaden, strengthen, etc.), whereas the indicators are states-of-being, starting with nouns and numbers (e.g. number of, % of, etc.). Indicators take shape as **absolute indicators** (e.g. such as 'number of deaths per 100,000') and **relative indicators** (proportions, rates and percentages representing how two or more absolute values relate to each other).¹²

10 ICSU, ISSC (2015). Review of Targets for The Sustainable Development Goals: The Science Perspective. <http://www.icsu.org/publications/reports-and-reviews/review-of-targets-for-the-sustainable-development-goals-the-science-perspective-2015/SDG-Report.pdf>

11 Bates-Eamer et. al (2012). Post-2015 Development Agenda: Goals, Targets, and Indicators. The Centre for International Governance Innovation and the Korea Development Institute. Available at: <https://sustainabledevelopment.un.org/content/documents/775cigi.pdf>

12 See Footnote 5.

Table 1: The Scheme of the Sustainable Development Goals Adapted from Bates-Eamer et al. (2012)¹³

	GOAL	TARGET	INDICATOR
Key definition	An ambitious commitment to an outcome addressing a single challenge	A specific, measurable and time-bound action directly contributing to the goal	An output metric used to measure progress towards the target
Data type	Qualitative or quantitative	Quantitative	Quantitative
Scope	Global	Global or national	Global or national
Example	16. Promote peaceful and inclusive societies for sustainable development at all levels	16.6 Develop effective, accountable and transparent institutions at all levels	16.6.2 Proportion of the population satisfied with their last experience of public services

CGD PROJECTS FOR LOCAL ACTION OFTEN USE DIFFERENT FRAMINGS THAN THE SDGS

As identified in a former report written by the authors for DataShift, the objectives of CGD projects vary greatly depending on the context and project setup.¹⁴ Most of the projects interviewed in the report had not yet considered using the SDGs as a framing for their work, beyond perhaps requesting funding from high-level donors. Some found that the focus on integrated sustainability went above their mandate of focusing on very specifically targeted local issues.¹⁵ However, CSOs and communities when educated directly about the SDGs through DataShift found the SDGs very relevant to their issues and found that they provide legitimacy to their work because of the recognition of the topic they focus on and placed them with a network of support of organisations working toward the same purpose. Also, if the SDGs are used as a framing from the beginning such as in Lanet Umoja, they can help guide community discussions around their own governance priorities.

13 See Suter (2014). Goals, Targets and Indicators. Definitions and Key Concepts For The Post-2015 Development Agenda. Independent Research Forum. Available at: http://www.irf2015.org/sites/default/files/publications/Retreat%20%232_Background_Paper_2_and_3_GTI_and_Criteria.pdf

14 For further information see, Lämmerhirt, D., Jameson, S. and Prasetyo, E. (2016). How to Make Citizen-Generated Data Work: Towards a Framework Strengthening Collaborations Between Citizens, Civil Society Organisations, And Others. Available at: <http://civicus.org/thedatashift/wp-content/uploads/2015/07/Making-citizen-generated-data-work.pdf>

15 Interview with Civic Action Group and Safecast, November 2016.

HOW CGD PROJECTS CURRENTLY LINK TO THE SDGS

In order to test how linking CGD to the SDGs would work in practice, a variety of CGD projects¹⁶ were matched as closely as possible to the SDG targets and indicators. To do so, we compared the unit of analysis of an SDG indicator with the measurement unit of related citizen-generated data. The table below presents a few examples from an exploratory mapping analysis of how CGD projects link to the SDGs and which indicator they best fit with. The exercise brings to light a **disconnect between CGD and the SDG indicators**, where the units of analysis are very different, making the connection very indirect.

Table 2: Examples of CGD initiatives mapped to SDG indicators

CGD INITIATIVE	DATA COLLECTED	SDG	TARGET	INDICATORS	HOW CAN CGD CONTRIBUTE TO SDG INDICATOR
Africa's Voices	SMS-based survey data to derive perceptions of sub-populations, combined with demographic data	16	16.6, 16.7	16.6.2; 16.7.1, 16.7.2	Provide rich qualitative context to understand the SDG action and monitoring process
Humanitarian OpenStreetMaps Team Indonesia	Mapping exposure to flood hazards	1; 11	1.5.1, 1.5.3; 11.3, 11.5, 11.b; 13.1, 13.3	11.3.1; 11.5.2, 11.5.2; 11.b.1; 11.b.2; 13.1.1; 13.1.2; 13.3.1; 13.3.2	Provide a baseline to measure outcome after disaster; provide context; data needs to be linked to action to create indicator data
Land Matrix Initiative	Land deals: Parcel size, countries involved, investment purpose	1; 12; 15	12.2, 12.6, 12.7; 15.1, 15.2, 15.3	1.4.2; 12.6.1, 12.7.1; 15.1.1, 15.1.2, 15.2.1, 15.3.1	Provide thematic overview on overlapping indicators; can create alternative indicators to highlight these issues
Lanet Umoja	Community awareness and advocacy on CGD for gender-SDG-related governance priorities	5; 16	5.1, 5.2, 5.4, 5.6, 5.a, 5.b	5.1.1, 5.2.1, 5.2.2, 5.4.1, 5.5.1, 5.5.2, 5.6.1, 6.6.2, 5.a.1, 5.b.1	SDG indicator is used as a guiding principle for creating CGD, so the data can feed directly into SDG framework.

¹⁶ For more information about these projects and the context within which they work, as well as other case studies in this report, see the previous report, footnote 13.

CGD INITIATIVE	DATA COLLECTED	SDG	TARGET	INDICATORS	HOW CAN CGD CONTRIBUTE TO SDG INDICATOR
Mobile Birth Registration	Biometric data registering children under 5	2; 3; 16;	2.2; 3.1, 3.2; 16.6, 16.9;	2.2.1, 2.2.2; 3.2.1, 3.2.2; 16.6.2, 16.9.1	If upscaled across entire country, can feed directly into indicator on neonatal mortality
Safecast	Geolocated granular radiation & air pollution measurements	3	3.9	3.9.1; 3.9.3	Causal link between baseline mapping & indicator on mortality rate attributed to pollution not clear; needs linking with national data on causes of mortality
WeFarm	Farmers sharing answers around agriculture & sale	2	2.3, 2.4	2.3.2, 2.4.1	Needs added incentive to collect data for the SDG as only secondary to purpose of project

CGD MOSTLY ADDRESS SDG TARGETS, NOT INDICATORS

Working on the SDGs explicitly means tracking progress over time against the specific targets and indicators. **CGD, on the other hand, is more instrumental**, driving or demanding change around a particular issue of importance to citizens, meaning it usually fits the targets rather than output indicators. For example, CGD projects measuring disaster risk and exposure, such as Humanitarian OpenStreetMap and Civic Action Group, directly address target 1.5: 'By 2030, build the resilience of the poor and [...] reduce their exposure and vulnerability to [...] shocks and disasters.' However, the indicators 1.5.1 and 1.5.2 measure this target via deaths and economic losses as a result of disasters. As preventative actions, the CGD projects do not measure the outcome of disasters but rather exposure as part of disaster risk reduction. It follows that there is a potential for CGD to support additional indicators where current indicators are limited, particularly on issues that can help the government fill data and capacity gaps, or where shadow monitoring is needed.

CGD NEEDS TO BE LINKED WITH OTHER DATA, USUALLY AT A NATIONAL LEVEL

The very process of creating CGD is dependent on ensuring citizens are incentivised to contribute data by making sure the project addresses an issue relevant for them, encourages the feeling of belonging to a community and being able to learn from others, and/or aligns with their personal values and interests.¹⁷

17 Lämmerhirt, Jameson and Prasetyo (2016): Making Citizen-Generated Data Work.

Available at: <http://civicus.org/thedatashift/wp-content/uploads/2015/07/Making-citizen-generated-data-work.pdf>

As a result, CGD projects may not have the incentive to create the data specifically for the SDGs, so other initiatives need to take it upon themselves to calculate the data linkages required to build up the SDG indicators.

However, these linkages can be methodologically challenging. On the one hand, many projects are local and are not replicated nationally, so would need to be linked with similar initiatives in order to have national coverage. Linking similar initiatives is easier on the local level than the national, because different communities have different issues and it's easier to connect to those more familiar with the local environment. Added to this, in order to fully measure the indicator the CGD may only deliver a part of the necessary data because their focus is on action. For example, the Mobile Birth Registration initiative¹⁸ can feed directly into the indicator on neonatal mortality, but would need to be combined with other vital statistics data to measure progress on children's health as they age.

Another challenge is that **some indicators imply a causal relationship that may not be captured in CGD**. For example, Safecast collects data on ambient radiation and air pollution measurements. The closest indicator this matches to is 3.9.1., '[Substantially reduce]' mortality rate attributed to household and ambient air pollution'. In this case, it is not enough to simply combine Safecast's data with mortality rates, because Safecast's data measures the state of pollution rather than causes of mortality. Instead, existing mortality rate causes from healthcare data could be triangulated with the Safecast data to reveal insights that are relevant for national and local action and decision making. This means CGD can deliver parts of the necessary data, or at least rich contextual information, whenever indicators measure a causal relationship.

Building on the example of Humanitarian OpenStreetMap above, if the preventive disaster risk reduction CGD projects want to link up to the SDG indicator, in principle, **the CGD could contribute to a baseline measurement** for the outcome indicators. The hazard risk mapping would be the 'before', and would need to be combined with survey data after a disaster to evaluate the damage. While in this specific example the outcome of disasters could also be mapped using CGD, this is not the stated goal of the CGD project as it currently stands—of course, they'd rather avoid that situation!

However, because CGD projects must be locally adapted in order to incentivise citizens to participate, it is not simply a matter of replicating projects across broader spatial scales to make sure there are more CGD projects to link together for a national indicator. Communities have different interests, and so the projects and CGD created will have different priorities as well. Linking these together is not

18 See <http://www.unicefstories.org/tag/birth-registration/>

a simple exercise of addition, but requires an analytical and political exercise of drawing out common interests to be able to aggregate the data. On top of these, concerns around donor and citizen fatigue, data security and implications of group privacy will also have to be handled as projects either take it upon themselves or are incentivised to create data linkages for the SDGs.

GOVERNMENT FRAMEWORKS AND STRATEGIES ARE DIFFICULT TO MONITOR AT SDG LEVEL

Some SDG indicators measure whether or not legal frameworks exist. For example, indicator 5.1.1: 'Whether or not legal frameworks are in place to promote, enforce and monitor equality and non-discrimination on the basis of sex'. In activist work, the presence of legal protection frameworks can indeed be a significant achievement for sustainable development. Similarly, other indicators focus on the outcome of operationalising and communicating policy around a particular strategic focus, such as indicator 13.2.1: 'Number of countries that have communicated the establishment or operationalisation of an integrated policy/strategy/plan which increases their ability to adapt to the adverse impacts of climate change [...]'.

However, beyond a successful campaign for lobbying for legal protection, CGD projects on those topics would consider the implication of the message inside those documents and how they are implemented as more important. As found in Lanet Umoja, the policy document is vital, but it is only a starting point to understanding the direction action based on CGD needs to take. This underscores the value of freedom of information requests and open data. Within the SDG framework of outcome indicators, it is difficult to measure these substantive monitoring questions.

DATA FOR THE SDG INDICATORS IS OFTEN TANGENTIAL TO CGD PROJECT AIMS

Depending on the CGD project strategy, collecting the data needed for the indicators may only be complementary, or even extraneous, to the goals of the CGD project. For example, WeFarm is a social network platform where small-scale farmers can exchange agricultural knowledge via SMS to increase their productivity and adaptability. As a result, WeFarm is instrumental to achieving target 2.3: 'By 2030, double the agricultural productivity and incomes of small-scale food producers [...] including through productive resources and inputs, knowledge [...]'. However, the indicators 2.3.1 and 2.3.2 measure volume of production and average income, which is information that is not actually required for WeFarm to achieve its aims of building the capacities of small-scale farmers.

Another example is Africa's Voices Foundation, who use radio and SMS to create vibrant civic discussions around development topics. The data is also analysed to help inform a variety of development programs. As a public forum for debate, this **CGD can provide rich contextual information** to interpret the SDG monitoring and action process. However, the CGD itself is about understanding public opinion, not monitoring progress over time.

These two examples shows a **disconnect between what is required of data:** there is a difference between the type of information needed to drive change on the ground, and the type of information necessary to understand and monitor developments at larger scale.

NOT ALL CGD ADDRESSES DEVELOPING COUNTRIES

The problem of coverage and representivity in CGD is well-known¹⁹, though usually in the context of local or national issues. Several CGD initiatives that work directly on the goals do not work with developing countries' data. For example, Patients Like Me (PLM) works directly with marginalised groups of patients with chronic diseases, and they make significant contributions to CGD for medical research. However, as a for-profit company, PLM's work is catered to its customer base, which is more from developed countries. The different customer base significantly influences the type of data collected; for example, whilst PLM does address HIV/AIDS patients, the concerns of an HIV-positive person in the United States who wants to track the effects of their retrovirals are very different from an HIV-positive person in Sub-Saharan Africa who may not even have access to the drugs.

In terms of connecting PLM's CGD with the SDGs, their work most closely aligns with target 3.3, ending the AIDS epidemic, and target 3.b, supporting the research and development of vaccines and medications. However, these targets focus specifically on healthcare in development, and particularly highlight medical issues 'that disproportionately affect developing countries'. Therefore, whilst PLM does excellent work for marginalised communities and their work could be linked to the SDG monitoring effort in certain countries, it does not directly align with the SDGs targets and indicators as they are currently written.

¹⁹ Datashift (2016). Making Use of Citizen-Generated Data.

Available at: <http://www.data4sdgs.org/guide-making-use-of-citizen-generated-data/>

COMPARING AND LINKING CGD

As seen above, CGD needs to be compared and linked with other data in order to link up to the SDGs. What follows is an overview of recommendations specific to the SDGs. It describes methodological issues around comparison, aggregation across spatial scales. For a further discussion, how what data means for action are further elaborated in the tandem piece, 'From Evidence to Action'.²⁰

As the aim is to make comparisons at a regional, national or global scale, the data or the sample needs to be statistically representative. Often this means that a broad coverage is necessary and the sample size needs to be large enough. For CGD, the challenge is to scale-up from projects responding to specific local contexts and then reproducing similar data collection methods in other areas. Let it be noted that this goes beyond a mere methodological challenge. If a CGD project wants to run a data collection in a city, local communities, political parties, administration, or business may all have different interests what to measure. The same problem arises if data shall be measured to compare across regions.

Conflicting interests can be aligned during the data design phase (for more information see tandem piece 'From Evidence to Action').

This phase not only serves to align political interests, but also to define reliable data collection methods; for instance **by standardising data capture methods and using data conventions** (for more detail and examples, see tandem piece 'From Evidence to Action'). Depending on the data these can enhance credibility of the data, facilitate interoperability between data sets, and foster common understanding between actors.

Questions of data coverage can be about spatial coverage for comparison across regions, but also temporal coverage for trend analyses. Trend analyses and SDG monitoring are particular in that they require data which is persistently captured over time. CGD projects that feed directly into e-governance monitoring processes aim to be long-term from the get-go. However, not all CGD projects aim for this longevity—for instance, some aim to address temporary problems.²¹

20 Available at: <http://civicus.org/thedatashift/learning-zone/>

21 As another paper of the authors discusses, responsible data management is key if transactional data such as dialogues are handled for analysis. Main concerns are how privacy may be safeguarded by anonymising and aggregating data adequately.

If this latter type of CGD is to be used for long term trend analyses or monitoring, there needs to be attention paid to how to do so without being burdensome and/or counterproductive. **Longitudinal snapshots can be a useful method to combine CGD with trend indicators** when ongoing real time monitoring is not feasible or productive, where the 'snapshot' that is captured by a short-term CGD project is then repeated (and thus, standardised) at a later stage (a few years later, for example), and the difference between the two states represents the trend.

Clearly, standardising across regions is much easier when a top-level actor working on that level can facilitate the data collection. Particularly in the context of the SDGs, this may very well be regional or national governments working towards the indicators anyway. However, effective collaboration around CGD is dependent on aligning interests,²² so thought needs to be paid to the strategy and who is initiating the project.

Standardisation is not a hammer looking for a nail, however. A major benefit of CGD is precisely the richness of local context and stories that it can provide. In the first phase of data collection, standardising the data (not the collection method) may be counterproductive, particularly with qualitative CGD aiming to show the diversity of experiences. Then, with careful attention and a standardised methodology in the second stage, it is possible to **combine quantitative and qualitative data for a more valid data set for the SDGs**. For example, in the historical case of the Ceará, Northeastern Brazil, ethnographic CGD provided complementary context and broader coverage for national statistics. By combining surveys, vital statistics and ethnographic research with 'popular Death Reporters', that is members of the community active in traditional death rituals, the state was able to use reduce underreporting with valid, qualitative mortality statistics.²³ The next section further explores question of representation and validity when data is aggregated for the SDG indicators.

22 Lämmerhirt, Jameson and Prasetyo (2016): Making Citizen-Generated Data Work.

Available at: <http://civicus.org/thedatashift/wp-content/uploads/2015/07/Making-citizen-generated-data-work.pdf>

23 Nations and Amaral (1991). Flesh, Blood, Souls, and Households: Cultural Validity in Mortality Inquiry. *Medical Anthropology Quarterly*, 5, 204–220.

AGGREGATING CGD FOR THE SDGS

From the point of view of CGD, data needs to be linked at larger scales and aggregated to fit the SDGs. From the point of view of the SDGs, national level indicators need to be “disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts”.²⁴ Therefore, no matter the starting point, it is crucial to understand the implications of aggregating data in order to think through how CGD can support the SDG monitoring effort.

Often local CGD is used to tell a story, and part of its strength is that it can offer stories and meaning that is complementary to official statistics, bringing issues to the fore that may have been missed. **CGD can significantly increase the validity of indicators**—the challenge is how to adequately represent marginalised voices throughout the different stages of aggregation that data goes through,²⁵ and keep that representivity as the SDG national indicator is disaggregated by population group. Aggregation is not necessarily a straightforward process.²⁶ What follows is an exploratory example thinking through how CGD that is complementary to official reporting could link up to the SDG indicators, at what stage, and what is necessary to do so.

AGGREGATING CGD FLOOD RISK DATA FOR THE SDG INDICATORS

Figure 1 shows the possible process of aggregating local CGD on flood risk and combining it with government data. Level 0 is the SDG indicator, and the diagram shows which data could contribute to it at what stage. The figure uses the data aggregation mode to match an empirical case to the SDG indicator 11.b and target 1.5. It draws from previous research²⁷ and interviews with the Civic Action Group in

24 United Nations (2015), Transforming our world: The 2030 Agenda for Sustainable Development. Available at: <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>

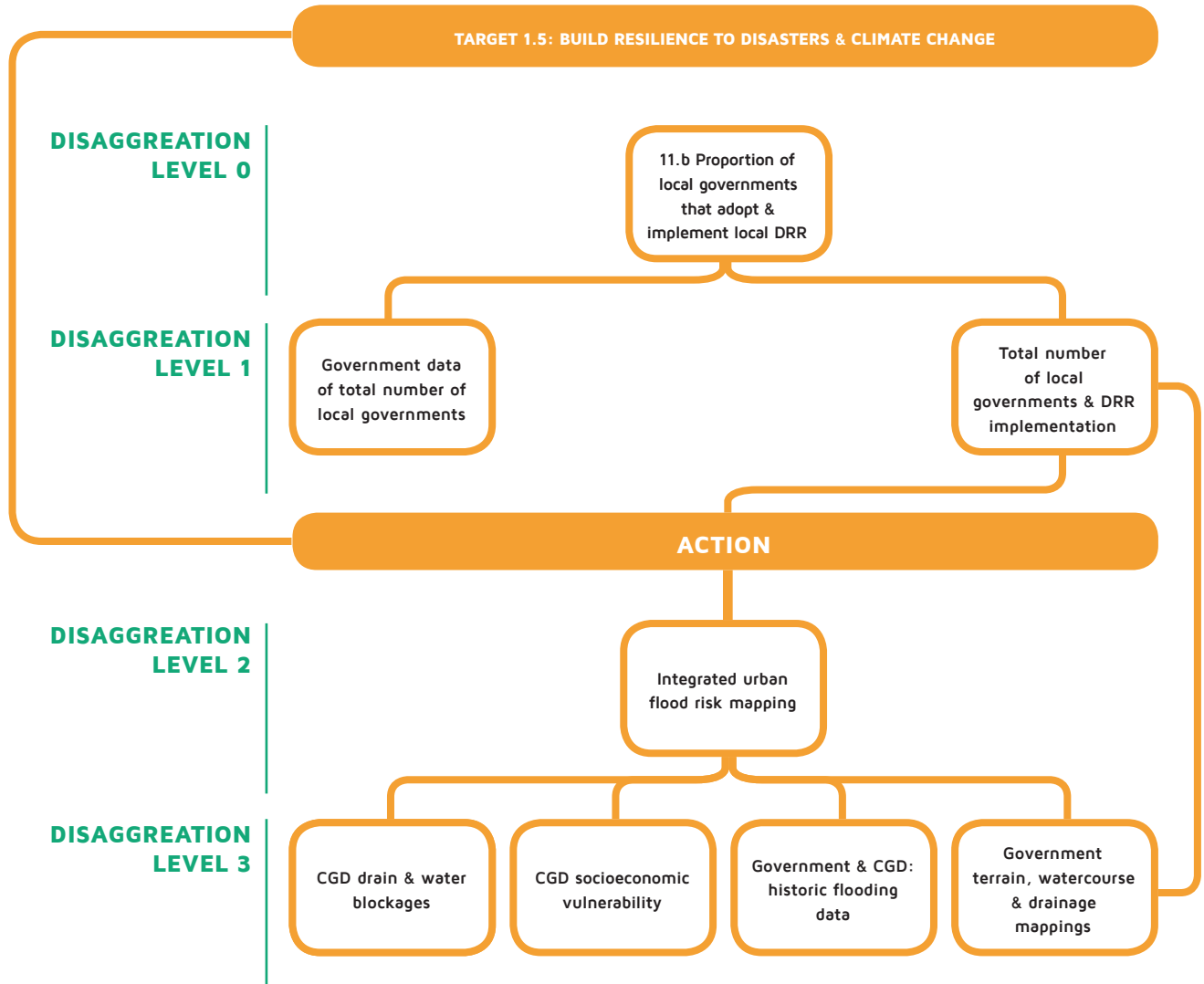
25 For example, as progressive as the SDGs are, they still do not have an operationalised answer to the classic feminist critique of traditional econometric indicators such as GDP, which states that these indicators are invalid because they do not include or represent unpaid household work, which is largely the contribution of women. Saunders and Dalziel (2016): Twenty-Five Years of Counting for Nothing: Waring’s Critique of National Accounts. Feminist Economics. Available here: <http://www.tandfonline.com/doi/citedby/10.1080/13545701.2016.1178854> Unpaid work is represented as a social protection, not an economic issue, in SDG indicator 5.4.1.

26 Once again, the specific methodological considerations are further elaborated in the tandem piece, From Evidence to Action.

27 Jameson and Baud (2016). Varieties of knowledge for assembling an urban flood management governance configuration in Chennai, India. Habitat International, 54, 112-123.

Chennai, India. Chennai flood management.org is an open data platform combining CGD from cartographic exercises along waterways and throughout the city. The platform geolocates different elements contributing to flooding (drains, blockages, encroachments, etc.), as well as official maps to create more integrated flood risk models. Civic Action Group plays a role in stimulating CGD mapping exercises (Level 3), as well as bringing data together (Level 2).

Figure 1: Hypothetical data aggregation model linking CGD vulnerability mapping to SDG indicators



The figure shows that in order to link to SDG indicators, the CGD needs to be combined with official data, both on the local level for more integrated mapping but also in order to move up spatial scales. Importantly, moving from Level 2 to Level 1 happens at another level of governance; urban level is interesting to the municipality and the local level, whereas an overview of the proportion of local governments as a whole is more interesting for the national level. In Chennai, coastal flooding is the mandate of the national government, whereas inland flooding is the mandate of the state and the municipal level, even though in some areas of the city the two types of flooding overlap in the same physical space. The different levels at which these issues are governed means it might make more sense to do comparisons across states or across cities, rather than regarding the nation as a whole. **The governance structure defines the necessary level of [spatial] data aggregation.**²⁸

What the figure also clearly shows is that the 'black box' of action is not only a key link to the SDG target, but also that fulfilling the criteria of the indicator 'implementing DRR strategies' is dependent on action on the part of the local government. Therefore, action is crucial, and in order to allow CGD to have impact and to link to the indicators, thought needs to be paid to **the important question of who is accountable in reporting and action**. From the CGD perspective, it needs to be clear who the data addresses and how to create a narrative that will best speak to them.



28 For further details, see also: Lämmerhirt, D., Jameson, S., Prasetyo, E. (forthcoming): From Evidence to Action.

RECOMMENDATIONS

Based on the report findings and exploratory mapping of different CGD projects to the SDGs, we recommend the following strategies for those seeking to enhance the SDG monitoring effort by leveraging CGD :

ENCOURAGE QUANTITATIVE MONITORING DATA TO BE PART OF CGD PROJECTS

CGD is often about filling data gaps by establishing a baseline of how things are in order to raise awareness or advocate for change on issues, rather than monitoring progress on government actions around those issue. SDG follow-up and review, on the other hand, is very much a monitoring effort, and only some of the indicators are about establishing baseline numbers. As seen in the Chennai flood mapping example above, CGD may not necessarily be able to complete the action required to address the framing of the SDG indicators, either for a lack of capacity or simply because it is the government's mandate. Therefore, if CGD is to be used for SDG monitoring, it is recommended to encourage quantitative monitoring data to become part of the CGD project initiative. More research is required to refine what kinds of incentives would facilitate this process and which actors could play a role here.

Within the variety of CGD project types, data can play a variety of roles. There are **three general approaches to the role of CGD within a project** (see Table 3): monitoring the extent of an issue, stimulating action, or providing a service through data. Naturally, some projects straddle different approaches, depending on whether the organisation initiating the project aims to stimulate action or not; some citizen science projects aim for neutrality by simply providing data, rather than engaging in political lobbying for action as well. The distinction between the different roles of data is critical because it highlights whether the type of data required to fit into the SDG framework, i.e. monitoring data around which actions have been taken, is inherently part of the data collection within CGD projects or not.

Table 3: How CGD can link to SDGs when data plays different roles within projects

GENERAL CGD PROJECT APPROACH	CGD EXAMPLE	HOW CAN THIS DATA CONTRIBUTE TO SDG MONITORING
Showing the extent / diversity of an issue	Citizen science projects, monitoring projects e.g. Safecast, World Water Monitoring Challenge	Use data to fill gaps in coverage and establish a baseline of how things currently are. If monitoring data is part of the project framing, this fits directly with the SDGs.
Stimulating action by presenting data	Local action e.g. Harassmap, FixMyStreet, Mobile Birth Registration	Use CGD project data on the result of a campaign / action can deliver evidence for progress around the SDGs.
Provide a service through sharing data	Knowledge platforms e.g. WeFarm, PatientsLikeMe	Create transactional CGD as a by-product that can be repurposed to monitor SDGs. ¹

WORK WITH AND BUILD ON EXISTING GOVERNMENT MONITORING PROCESSES

A lot of the data required for the SDG indicators is government data, such as descriptive population statistics, level of government mandated for action, and surveys such as causes of mortality or land registries. This means that in order to combine CGD with government data, the government data has to be accessible (preferably open), but also the ways in which data is dealt with within government need to be understood. Because action on the issue at hand is central, and the recurring question is always what helps or hinders the uptake of CGD, **SDG monitoring projects need to talk to governments** in order to see how they can be of most support to existing monitoring processes. Building on existing processes also helps link to the SDGs without the SDGs feeling overly like an additional administrative burden.

USE SDGS AS A COMMON FRAMING

As an accountability mechanism, SDG monitoring is meant to incentivise integrated action for sustainable development across parties within a particular framework. Collaboration is vital, and the politics of discussion and navigating agendas between parties is not always discussed in high-level monitoring documents. Previous research, as well as the tandem piece to this paper, has also underlined the importance of aligning interests in order to make CGD projects effective in their strive for change and encouraging the uptake of CGD. **The SDGs can help provide a common language** to move towards integrated development because of the international standing of the framework.²⁹

In the direct support work that DataShift has been doing with CGD initiatives around the SDGs, CGD projects have seen the benefits of the SDGs as a means of providing legitimacy to their issue as it is already recognised in UN resolutions. **The SDG framework can also serve as a linking mechanism** to connect projects working on similar issues within and across countries, which can benefit both the sharing of best practices, data conventions, as well as networking and encouraging collective action. This common framing can also encourage different CGD projects to connect their data for more integrated measurement of the SDG indicators.

Help build alternative indicators with CGD that represents cross-cutting issues

Some projects work on specific topics at the intersection of several SDG targets. For example, Land Matrix Initiative gathers data on purchasing party, land size and or parcel location. It thus directly addresses targets on land rights, sustainable procurement practices, and transparency and accountability (Goals 1, 12 and 16). Because large-scale land acquisitions in developing countries are often linked to extractive industries and monoculture cropping, and Land Matrix's database

29 For further details, see also: Lämmerhirt, D., Jameson, S., Prasetyo, E. (forthcoming): From Evidence to Action.

records the purpose of the investments³⁰, the issues of sustainable agriculture and ecosystem conservation are important to connect to (Goals 2 and 15). This is a challenge and opportunity at the same time: **CGD may be potentially deliver data that can be more flexibly applied across different targets**, thereby helping to facilitate overcoming silo thinking when used for different analyses that are based on a common denominator like land purchases. While the data of Land Matrix Initiative do not reflect official indicators they offer material for alternative indicators including those related to transparency, such as ‘percentage of land acquisition deals that are well-documented in open data registers’.

BUILD LOCALLY-RELEVANT INDICATORS AROUND SDG 16 WITH CGD

The indicators for this goal are all national level indicators, so local-level CGD projects have the room to contribute locally-relevant indicators within this framework, as shown in figure 1 (linking vulnerability mapping to the SDGs). A key role for several CGD projects, particularly those directly seeking to engage the government, is to hold governments accountable and help build institutional capacities. The social-auditing model is an important and well-known mechanism for CGD projects, such as mySociety’s Fix My Street. As a result, several CGD programmes have a lot of impact on Goal 16, working towards ‘effective, accountable and inclusive institutions’. Because institutional capacity is hard to define, **the targets under SDG 16 are also some of the most open-ended** in the official set of indicators, and specifically address ‘all levels’ of governance. This leaves a lot of room for CGD projects, particularly those working on local-level issues, to slot into the SDG framework. At a broader scale, one example of an alternative indicator around institutional capacity is the Civicus Monitor,³¹ which tracks civic space by triangulating several data streams on democracy indexes, media monitoring, and research, amongst others.

BUILD METHODOLOGICAL CAPACITIES AROUND STANDARDISATION

Standardisation is necessary for methodological rigour in order to compare and build the indicators, both in terms of the type of data collected, data collection methodologies and data processing methodologies. This includes how qualitative data stories are converted into quantitative information for more valid indicators. Even the UN Independent Expert Advisory Group has acknowledged the need to standardise working streams for SDG monitoring.³² Importantly, many local CSOs, CGD projects and even governments do not have the technical capacity for

30 http://landmatrix.org/media/filer_public/ab/c8/abc8b563-9d74-4a47-9548-cb59e4809b4e/land_matrix_2016_analytical_report_draft_ii.pdf

31 See also <https://monitor.civicus.org/>

32 See UN data revolution: <http://www.undatarevolution.org/report/>

standardisation across the board. Therefore much of the investment required in the SDG monitoring effort, from both within and outside government, is to build out these capacities.

Ideally, there would be **standardised methods and data conventions for each type of indicator**. At a global level this may help national governments frame their monitoring efforts. At a local level, data conventions between parties can help build the capacity of citizens and their organisations to contribute to SDG monitoring. One example of this is the ongoing DataShift initiative collaborating with the Tanzanian Bureau of Statistics to create a handbook of guidelines for civil society on how to create reliable, useable data which complies with the Tanzanian Statistics Act of 2015.

LINKING CGD WITH THE SDGS—WHICH ARE THE COLLABORATIONS AND GOVERNANCE ARRANGEMENTS REQUIRED?

CGD projects are designed to solve an issue by addressing and engaging with stakeholders. Only if CGD is relevant, it can steer behaviour, or enable decision-making of those stakeholders. **The question is paramount, for whom is which type of information useful, when trying to scale locally collected data to regional and national levels.** It is important to recognise that **the governance around an issue defines the necessary level of [spatial] data aggregation required**. For instance: National government bodies may be responsible to allocate money to regions for water point constructions. Responsibility for their maintenance may reside with local districts. While national government needs comparative data across regions to allocate infrastructure investments, local districts need hyperlocal waterpoint information. An ongoing monitoring has to be conducted on a district-by-district basis, built on hyperlocal maintenance data. Before intending to scale CGD, it should be asked **for whom does data at another scale matter**.

Linking CGD to the SDG monitoring effort will require collaboration across the board, though certain actors are better placed to address some of the points highlighted in this report. For all initiatives seeking to build these linkages, be they government or civil society, mapping the governance levels involved as data moves up scales is crucial. Using the SDGs as a common framing for collaborations and building capacities for methodological rigour and standardisation will also be important for civil society and government collaborations. Some aspects, such as creating alternative indicators and issues around civic space in SDG 16, will be better taken up by civil society holding governments to account.



Shazade Jameson, Danny Lämmerhirt,
Eko Prasetyo, *Acting locally,
monitoring globally? How to link
citizen-generated data to SDG
monitoring.*

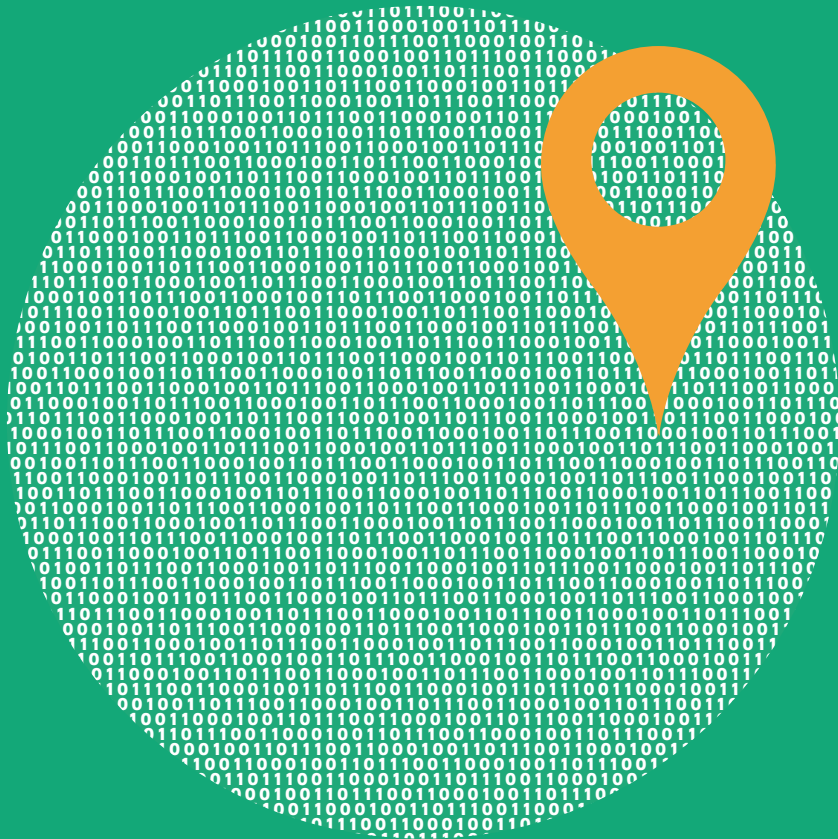
For more information, visit www.thedatashift.org
or contact datashift@civicus.org
First published, January 2017



This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Edited by Jack Cornforth and Hannah Wheatley
Printed on recycled paper/FSC.
Graphic design by Federico Pinci





Join the **DataShift Community** of civil society organisations, campaigners and citizen-generated data and technology practitioners by signing up at www.thedatashift.org and follow us on Twitter [@SDGDatashift](https://twitter.com/SDGDatashift)

DataShift is an initiative of **CIVICUS**, in partnership with **Wingu**, **The Engine Room** and the **Open Institute**. We are part of a growing global community of campaigners, researchers and technology experts that is using citizen-generated data to create change.

