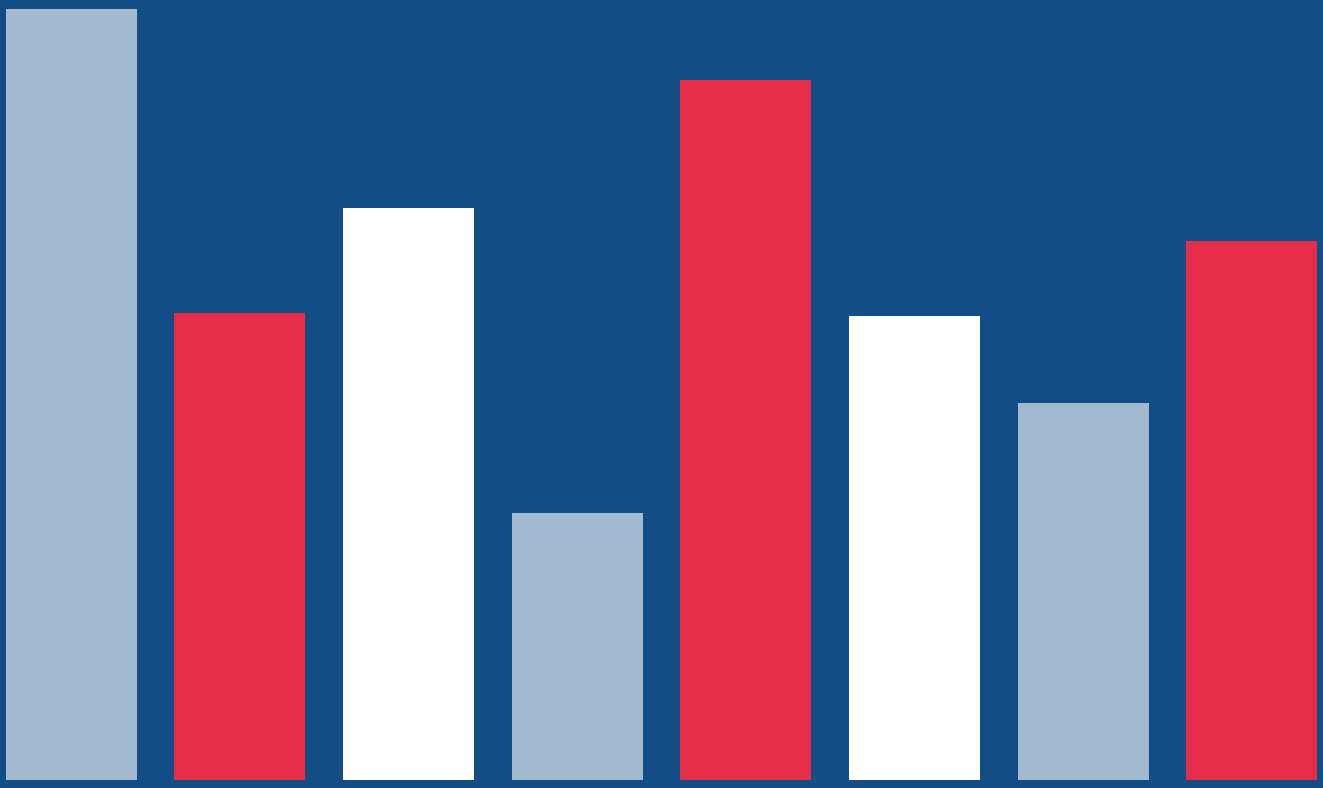
STATISTICAL PERSPECTIVES ON CITIZEN GENERATED DATA

Federico Piovesan

DATASHIFT



Federico Piovesan
is pursuing a PhD
in Urban and Regional Development
at the Polytechnic of Turin:
his research focuses on
the use of civic technology
for participatory practices.

STATISTICAL PERSPECTIVES ON CITIZEN GENERATED DATA

Federico Piovesan

ACKNOWLEDGEMENTS

We would like to express our gratitude to all those we interviewed for their valuable input. The following list includes respondents who agreed to have their name and organisation published.

- Natalia Aquilino, Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento (CIPPEC)
- Rajan M. Bajracharya, International Center for Integrated Mountain Development (ICIMOD)
- Krishna Hari Baskota, Nepal National Information Commission
- Mitch Blaser, MCC-PEPFAR Data Collaboratives for Local Impact
- Rudi Borrmann, Modernización y Gobierno Abierto
- Stephen Chacha, Africa Philanthropic Foundation
- Carolina Cornejo and Olivia Minatta, Asociación Civil por la Igualdad y la Justicia (ACIJ)
- Agustina De Luca, Fundación Directorio Legislativo
- Al Kags, Open Institute
- Linet Kwamboka, Kenya Open Data Initiative
- Daya Sagar Shrestha, NGO Federation of Nepal
- Karen Rono-Bett, Development Initiatives

TABLE OF CONTENTS

1 BACKGROUND 7

2 METHODOLOGY AND RESULTS 9
- 2.1 Country focus: Argentina 11
- 2.2 Country focus: Kenya 14
- 2.3 Country focus: Nepal 17
- 2.4 Country focus: Tanzania 20
- 2.5 International non-governmental organisations 22

3 DISCUSSION AND OPPORTUNITIES 24
- 3.1 Argentina 25
- 3.2 Kenya 27
- 3.3 Nepal 28
- 3.4 Tanzania 29

4 IMPLICATIONS FOR NATIONAL MONITORING STRATEGIES 31
- 4.1 Argentina 31
- 4.2 Kenya 32
- 4.3 Nepal 33
- 4.4 Tanzania 35
- 4.5 International non-governmental organisations 36

5 CONCLUSIONS 37

6 RECOMMENDATIONS FOR DATASHIFT AND PARTNERS 40

REFERENCES AND FURTHER READING 42

ANNEX1 INTERVIEW QUESTIONS 44

WHAT IS CITIZEN-GENERATED DATA?

Citizen-generated data (CGD) is data that people or their organisations produce to directly monitor, demand or drive change on issues that affect them. It is actively given by citizens, providing direct representations of their perspectives and an alternative to datasets collected by governments or international institutions.

Increasing numbers of CGD initiatives across the globe are using various methods and technologies to collect data for research, awareness raising, and advocacy. Examples of how this data is collected range from traditional paper-based questionnaires, SMS, and radio broadcasts to geo-referenced data collection (including text and media) via smartphones.

HOW TO READ BOX PLOTS

Box plots on the following pages summarise the views of respondents about the opportunities for CGD to contribute to SDG monitoring.[1]

Box plots are used to visualise the distribution of the answers in each group. The top and bottom of the box represent the first and third quartiles, respectively, while the line within the box represents the median observation (which corresponds to the second quartile). Finally, the whiskers (the lines extending vertically from the box) represent the maximum and minimum observations.[2]

Boxes provide a more complete picture of the average of an option's overall rank. Boxes that are small and focused in a given position show agreement on whether an option has low or high priority. Boxes that span through the height of the graph, on the other hand, mean some participants considered that option as high priority whereas others did not.

1 This publication includes only the graphs that summarise results for question 4 (see annex 1 for the full questionnaire). A subsequent version will include all the graphs in an additional annex

2 Box plots are non-parametric, meaning their strength cannot be tested against any standard statistical distribution.

BACKGROUND

The process to create a new set of global Sustainable Development Goals (SDGs) has been accompanied by an increased focus on data as a key factor in both driving and measuring progress on sustainable development. According to the UN Independent Expert Advisory Group on the Post-2015 Agenda – one of the first groups to call for a "Data Revolution for Sustainable Development" – data "is a key ingredient in creating more mutually accountable and participatory structures to monitor the new goals".

The Millennium Development Goals (MDGs) demonstrated the limits of a top-down prescriptive approach to achieving progress on sustainable development, especially in underrepresented communities where opportunities and tools for participation, advocacy and accountability are lacking.

However, modern information and communication technologies (ICTs) present us with opportunities to collect more data, more frequently, including in remote areas, while facilitating replicability and containing costs.

CGD can be used to confirm or question official statistics and the accuracy of the stories and trends they convey. However, there is also much potential to actually integrate CGD within government datasets used to shape policy and report to the public or international bodies, so as to be able to provide a more detailed and up-to-date account of progress.

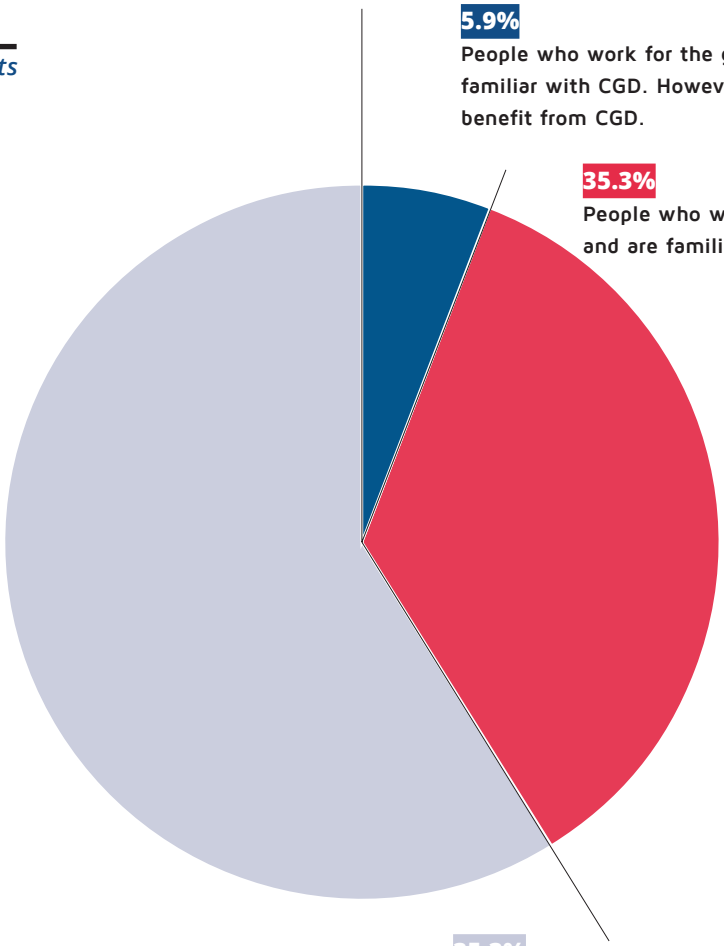
Moreover, by channelling voluntary contributions of data and information from individuals on issues relevant to them and their communities, CGD initiatives can help policymakers address gaps in their knowledge and understanding about the diverse needs of their constituents – especially in those communities that tend to be excluded from conventional participation processes. CGD can thus help crowdsource data collection for sustainable development that truly "leaves no one behind".

However, despite its promising potential, there are currently few examples of CGD being integrated into official datasets.

While civil society can be seen as a trusted producer of data to nuance policy, this is not always the case. Furthermore, involving citizens (who often possess limited statistical training and knowledge about political processes) in data collection presents new problems around credibility in particular. Through this study, therefore, DataShift has sought to work with official actors within the official statistics community to unpack these issues and identify both challenges and opportunities for integrating CGD into government datasets.

FIGURE 1

Participants



5.9%
People who work for the government but are not familiar with CGD. However, their work can potentially benefit from CGD.

35.3%
People who work within government and are familiar with CGD.

35.3%
People who work within government and are familiar with CGD.

2 METHODOLOGY AND RESULTS

We conducted 17 interviews with a selected group of participants, divided into 3 groups as shown in *Figure 1*. Key characteristics included the participant's profession (whether public employee or non-governmental actor) and his/her familiarity with CGD.

We first identified a number of barriers to the integration of official and CGD based on our experience as supporters and promoters of CGD initiatives (summarised as *Challenges* in the box below). Additionally, we surveyed interviewees' perceptions about the ways governments can benefit most from CGD and, consequently, the actions to which public and civic actors should devote their efforts in the short term (*Opportunities* in the box below).

Three-quarters of participants (76.5%) were familiar with CGD but less than half (41%) had direct experience working with CGD or organisations that produce it. CGD is a relatively new concept – at least in the terms previously discussed. However, some of the interviews highlighted how there was a need for a more consolidated definition, as CGD may still mean different things to different people.

This section delves into the perceptions of interviewees with respect to CGD, the opportunities it brings, and the challenges it currently faces. Section 2 then provides a summary of findings from each DataShift pilot country, along with actors working at an international level. Subsequently, section 3 expands on the challenges and opportunities around using CGD in these national contexts by combining findings from the interviews with previous research. Finally, Section 4 evaluates the observed attitude of officials and focuses on the short-term actions that can facilitate dialogues to promote inclusive monitoring.

CHALLENGES

- Unfamiliarity of statisticians and policymakers in government with CGD, which to them may represent a disruption of traditional data collection.
- Lack of endorsement from official institutions, thus preventing CGD from having a significative impact on policy.
- Technical issues, including:

  Lack of representivity: data represents only a limited group of people and thus is not representative of a larger group or geographic area.

  Lack of methodological rigour: officials see CGD's quality as unreliable because of its diverse collection and verification methods.

  Complementarity: the issues monitored through CGD are not always comparable or complementary to those analysed by governmental agencies.

  Interoperability: the format of CGD is such that it requires further processing before it is compatible with official data.

  Coverage: data covers a geographical area that is not comparable to the area analysed by official observations (too narrow or too wide).

  Persistency: there is no assurance an initiative will continue to operate and data will continue being collected in the future.

  Weak or adversarial relationship between civil society and national statistical and monitoring professionals.

OPPORTUNITIES

- Using CGD to complement or bolster official data in monitoring and spurring action on the SDGs.
- Sectors where governments can benefit most from CGD by gaining direct insights from citizens (e.g. education, health, gender equality, etc.).
- Approaches to promote the pluralistic monitoring of sustainable development in the shorter term, such as:

  Nurturing collaboration between official and civil society data producers:

  Collaborative data catalouges between governments and civil society.

  Multi-stakeholder workshops to jointly develop data collection and use guidelines.

  Development of common standards.

  Secondments and fellowships between civil society and governmental agencies to strengthen dialogue and collaboration around monitoring.

  Collaborative monitoring that goes beyond the integration of separate CGD sets within official portals.

  Capacity building programmes to support data collection and monitoring within civil society.

  Integrating CGD into government datasets:

  Investment in further data collection operations.

  Adoption of the proposed data collection practices.

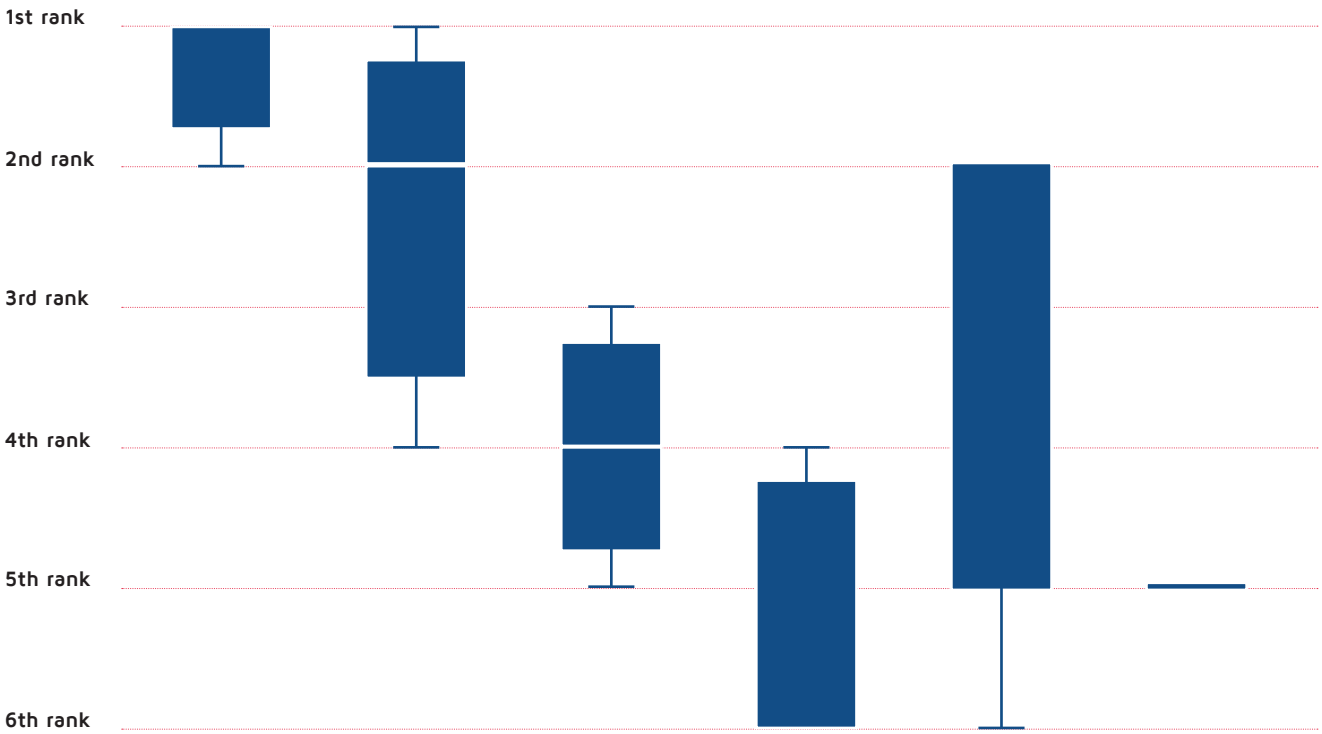
  Engagement with CGD collection.

  Official endorsement and recognition of CGD collection.

2.1 COUNTRY FOCUS: ARGENTINA

ARGENTINA: QUESTION 4

"How could citizen-generated data contribute best to monitoring the Sustainable Development Goals (SDG)?"



Answers

Validating or questioning official statistics.	Identifying areas for further data collection.	Increasing data coverage.	Identifying potential programmes and policy responses.	Additional information to complement official indicators.	Providing local level context.

Opportunities.

Respondents from Argentina agreed CGD could be useful to *validate or question official statistics*. Indeed, "sometimes public officers are too in love with their programme and cannot see whether it is actually impacting people's lives", according to N. Aquilino (CIPPEC) when discussing the need for alternative views to official data. It is interesting to see, however, that *providing local-level context* was consistently ranked second to last.

There was not as much consensus on other aspects. For example, both *identifying areas for further data collection* and *additional information to complement official indicators* received mixed rankings. This may indicate that CGD is seen more as ancillary to official data rather than as a source for unconventional perspectives on social issues.

Respondents also proposed some venues for CGD to contribute to SDG monitoring. First, CGD can promote more inclusive dialogue about monitoring and implementing the Goals, thus expanding the discussion to new non-governmental players (whether these are data producers or not). Moreover, CGD initiatives can help diffuse data culture and awareness among people, for example on topics like digital footprints on social platforms.[3]

Sectors.

Some respondents have had experience with CGD projects on specific issues, including housing and urban development (e.g. identifying slums[4]), monitoring of public service delivery, health care, and collecting and diffusing information about legislators.[5] However, most agreed that, although these examples prove CGD's potential to contribute to individual goals, opportunities to include citizens in data collection should not be confined to a limited set of areas.

For instance, R. Borrmann (Ministerio de Modernización y Gobierno Abierto) agreed that "the classics are important" (referring to data on census, health, education, agriculture, etc.) but governments also need to start considering (and preparing for) real-time data,[6] "[which] represents a whole new level of information with lots of potential and, though the government is currently unable to exploit it, it is

3 Digital footprints are traces that users produce through interaction with online platforms. While active digital footprints are data that users deliberately share, passive ones are collected without notice. Most websites and smartphones apps collect data about usage, and many base their business model on the sale of this information for marketing purposes. For more information see **http://www.pewinternet.org/files/old-media/Files/Reports/2007/PIP_Digital_Footprints.pdf.pdf**

4 For the project's page see **https://www.caminosdelavilla.org**/

5 See, for example, **http://directoriolegislativo.org/directorio**/

6 Real-time data (also called hot or dynamic data) is data that comes from electronic devices (like sensors) that upload each observation to a server for storage and reuse in a variety of applications. Think, for instance, of public transit data or environmental monitoring devices.

paramount that real-time data gets acknowledged and policymakers start planning how to use it". Partnerships with companies that produce low-cost sensors may bridge the lack of real-time data expertise, while also proposing innovative business models that can help CGD initiatives fund their activities.

A. De Luca (Fundación Directorio Legislativo) pointed out that "it is more about rigour and data accuracy as well as recognition and legitimisation from public figures" rather than specific areas of application. Citizens need training to collect data that can actually have an impact on policymaking but, at the same time, they must feel officials will consider CGD seriously within their decision making processes.

Challenges.

Participants agreed on the importance of CGD as a source of information about very localised (and often underrepresented) contexts. Hence it is important that CGD can be replicated across wider areas. Additionally, some participants mentioned that citizens – and often civil society organisations (CSOs) lacked the material resources (mainly financial) to replicate governmental data collection. Although ICTs can facilitate collection at a large scale, the limit nature of resources may affect data quality.

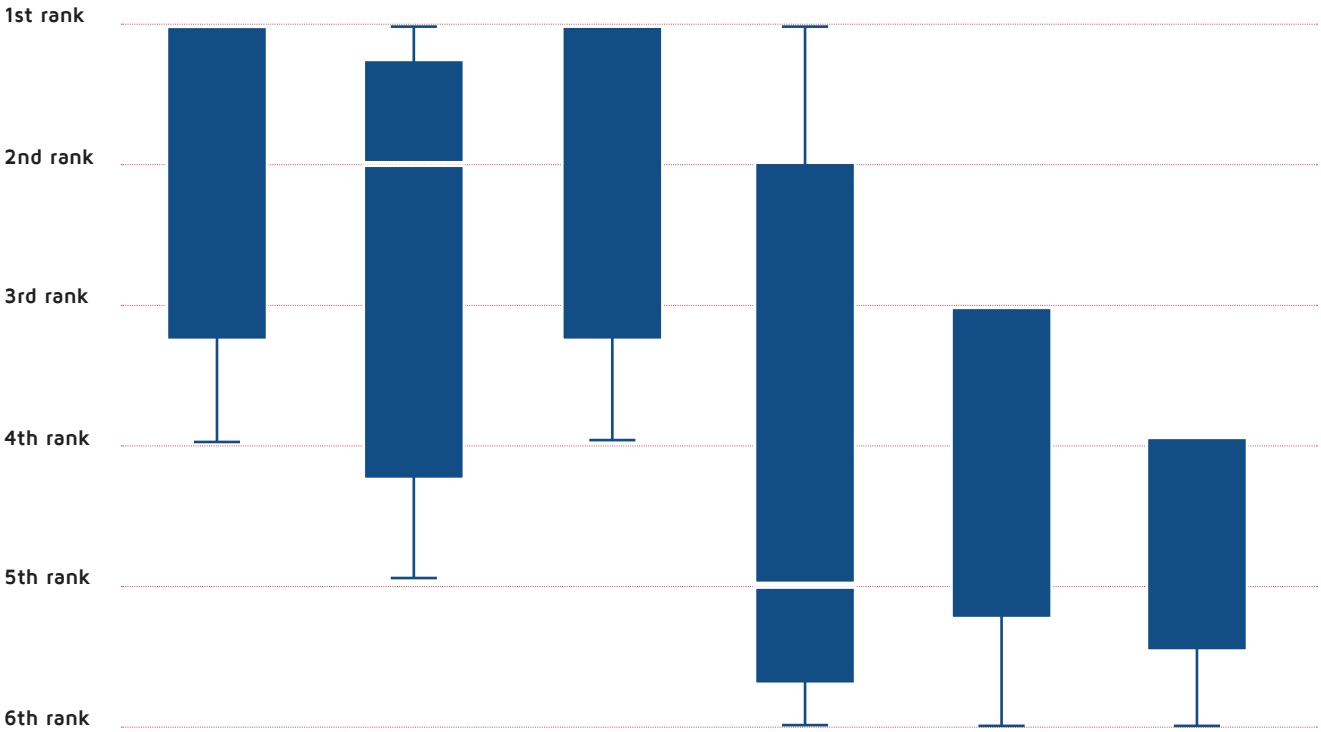
While opinions differed on *interoperability* and the *relationship between civic and official data producers*, some respondents prioritised the need for citizens to receive training directly from public statisticians. Indeed, most participants ranked *lack of methodological rigour* and *complementarity* (which implies understanding of official monitoring processes) highly.

Finally, limited financial resources emerged as an obstacle to the success of CGD initiatives. N. Aquilino's organisation (CIPPEC), for instance, has been planning a CGD project, but this cannot start because of a lack of funding. She also mentioned that Argentinian CSOs needed to improve their relationship with international organisations that can fund their projects.

2.2 COUNTRY FOCUS: KENYA

KENYA: QUESTION 4

"How could citizen-generated data contribute best to monitoring the Sustainable Development Goals (SDG)?"



Answers

Validating or questioning official statistics.	Identifying areas for further data collection.	Increasing data coverage.	Identifying potential programmes and policy responses.	Additional information to complement official indicators.	Providing local level context.

Opportunities.

Validating or questioning official statistics and *increasing data coverage in contexts where resource constraints or other considerations limit available data* had the highest aggregate scores, though agreement was not unanimous. *Providing local-level context*, on the other hand, received mostly low rankings. This may suggest that interviewees from Kenya are concerned mainly with the data gaps in official statistics but lack of apparent consensus on most options does not provide robust support to this idea. Rather, it shows that different actors have different ideas about how CGD can help in monitoring the SDGs.

"Governments do not necessarily need to collect data every five or ten years; there are simpler and cheaper ways to do it now", said K. Rono (Development Initiatives), who thinks CGD should be promoted as a tool for both advocacy and administrative purposes. "First", however, "we need to create a space where CGD is considered accurate and credible" and, while official endorsement is necessary, CSOs should help define what is and what is not considered CGD. "We should expand our definition, or sharpen it, depending on the context", she added, while highlighting how concrete examples were necessary to make the idea stick in people's mind.

Sectors.

Respondents mentioned census, development, health, and data about budgetary processes. None of them referred to specific projects, though examples of CGD initiatives in Kenya include data on education (like the School Report Card Project and Not in My Country),[7] health (Older Citizens Monitoring),[8] and social and development issues (On Our Radar, Development Check, and the Map Kiberia project).[9]

Most interviewees thought there were promising starting points for CGD to be a good complement to official data in many areas. Moreover, "CGD can help identify the best ways to communicate current results on sustainable development progress and how to disseminate these to wider audiences", said L. Kwamboka (Kenya Open Data Initiative). Hence CGD should not only ease a government's monitoring duties but also be useful to the citizens who produced it.

Challenges.

All participants considered *persistency* a very important issue: A. Kags (Open Institute) suggested that "we should not underestimate fatigue within communities": citizens need to see that the data they produce achieves some impact, since funding may not be the only determinant of a project's sustainability.

7 See **http://devinit.org/#!/post/using-citizen-generated-data-to-improve-school-performance-in-kenya**

8 See **http://www.helpage.org/silo/files/ocm-guidelines.pdf**

9 See, respectively, **http://www.onourradar.org/kenya/about-the-project/**, **http://www.developmentcheck.org/** , and **http://www.mapkibera.org/**

Methodological rigour, *complementarity*, and *coverage* scored consistently as either "somewhat important" or "very important" issues. "We need to demonstrate what gap CGD is filling", said K. Rono (Development Initiatives) on *complementarity,* so that CGD can gain credibility and address the official statistics audience. Especially when projects aim at supporting official monitoring "organisations need to have a clear channel of communication with the relevant office for the data they are providing", according to L. Kwamboka (Kenya Open Data Initiative). Therefore if this type of CGD does not reach the right officials, its impact will likely remain fairly limited.

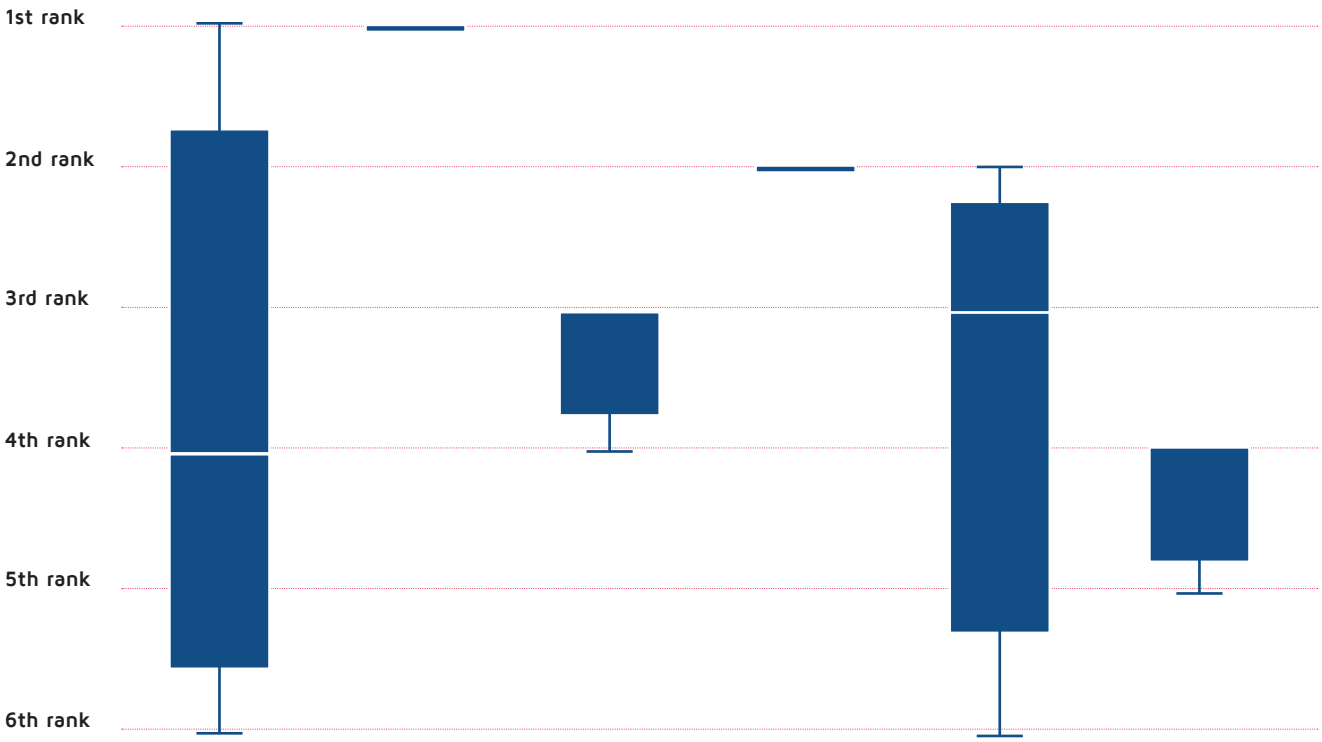
"In locations where there are no CSOs there is probably no data too", added K. Rono (Development Initiatives) on *coverage*, stressing the need for standardised data collection that can be spread to "make sure that no area gets ignored while others are saturated". She added, however, that standards should provide a solid foundation while staying flexible enough to adapt to the specificity of each context: "that's the beauty of CGD, it doesn't need to be a blanket that fits everyone, otherwise it would lose its citizen flavour."

While *the nature of relationships between civil society and national statistical and monitoring professionals* received mixed rankings, K. Rono (Development Initiatives) described how this could not be underestimated: "officials do not trust CGD and civil society generated data because they do not trust officials. We need the trust first and foremost."

2.3 COUNTRY FOCUS: NEPAL

NEPAL: QUESTION 4

"How could citizen-generated data contribute best to monitoring the Sustainable Development Goals (SDG)?"



Answers

Validating or questioning official statistics.	Identifying areas for further data collection.	Increasing data coverage.	Identifying potential programmes and policy responses.	Additional information to complement official indicators.	Providing local level context.

Opportunities.

Nepali respondents had mixed familiarity with CGD: as a matter of fact, only one of all organisations examined had written plans to engage with crowdsourced data collection. However, everyone expressed interest in CGD and its potential to assist the government in SDG monitoring. Answers suggested (some more subtly, some more explicitly) that official agencies, however committed to sustainable development, did not have the means to collect all relevant data on their own.

Indeed, the previous graph shows how all participants ranked *identifying areas for further data collection* as the most important opportunity. *Identifying potential programmes and policy responses* and *increasing data coverage*, on the other hand, were in second and third position, respectively. There was no clear consensus *on validating or questioning official statistics*. While some participants deemed it interesting, others considered it a troublesome or unrealistic prospect given that CGD is still unknown and will not be accepted easily in official circles.

Sectors.

Participants agreed CGD could span across a wide range of issues. R. M. Bajracharya (ICIMOD) noted that there were more than 60,000 community-based organisations in Nepal; providing them with an accessible and standardised framework for data collection would foster CGD initiatives that cover all the SDGs and beyond.

Respondents did not mention any individual projects, though they pointed to health, education, and natural resources as particularly promising sectors for CGD in Nepal. The Dalit Welfare Organisation is one of our network's initiatives that deals with (among others) health and education issues.[10]

"Nepal abounds with natural riches but its government holds little information about it", said K. H. Baskota (National Information Commission), who felt CGD could contribute to filling this data gap. The Community and Self Reliance Center, for example, addresses land and agrarian rights in rural Nepal.[11] The Open Development Network and Resilience Atlas, on the other hand, already offer interactive cartographies about the presence and use of natural resources in different countries.[12] Their platforms are based on open data and are designed to include crowdsourced information, and thus may offer useful insights and tools to launch similar projects in Nepal.

10 See **http://www.dwo.org.np/**

11 See **http://www.csrcnepal.org/**

12 Whereas the first focuses on Southeast Asian countries the second covers most of the world. For more information see **https://opendevelopmentmekong.net/** and **http://www.resilienceatlas.org/**, respectively.

Challenges.

Opinions about challenges are in line with those on opportunities: *complementarity* and *coverage* were seen as the most important issues, as CGD should support official data. Other issues, considered less important but still relevant, were *lack of methodological rigour*, *persistency*, and the *relationship between government and civil society*.

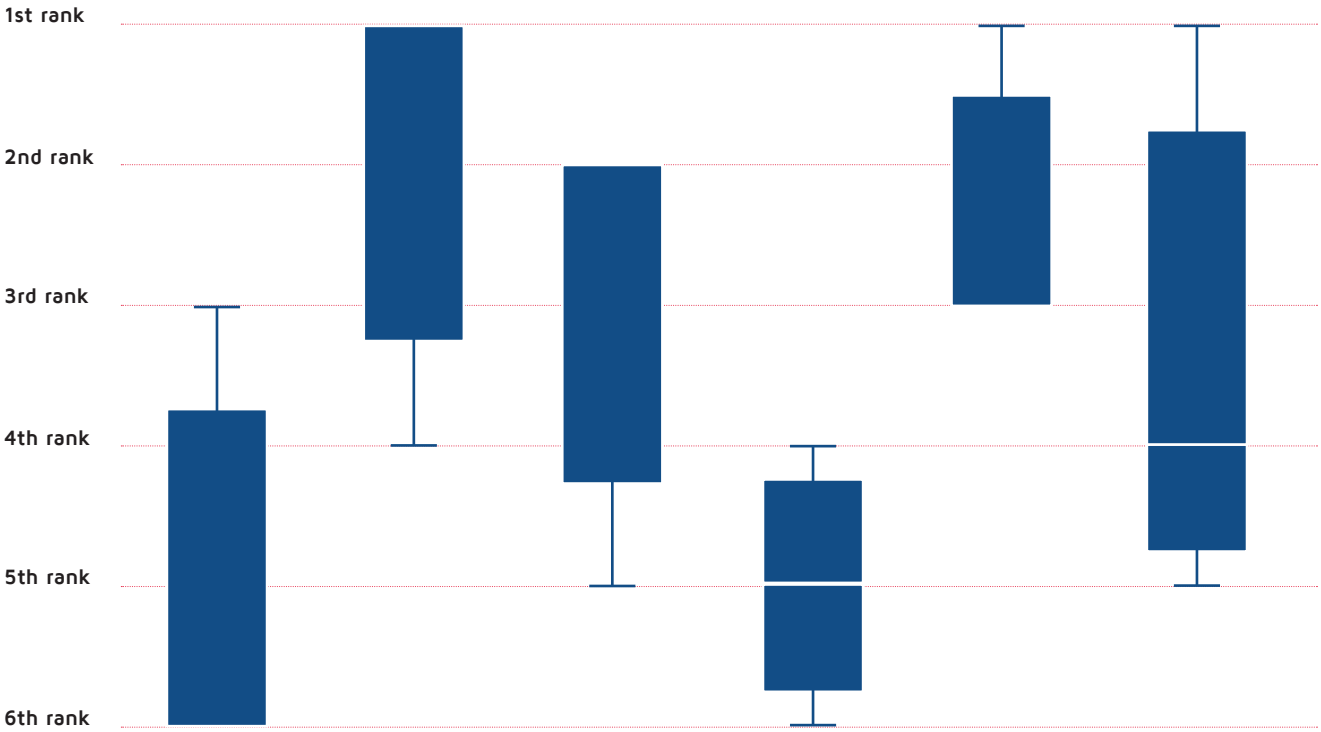
K. H. Baskota (National Information Commission) showed concern about the *persistency* of CGD initiatives, citing a lack of "resource, passion and courage." A R. M. Bajracharya (ICIMOD), on the other hand, was more interested in the need for a common framework regulating multi-stakeholder monitoring of each sustainable development area.

He also argued that both national and international institutions should acknowledge how time- and resource-consuming conventional data collection is: "the UN and the government could assist local initiatives that use innovative approaches to fill data gaps. For example, satellite imagery can support mapping initiatives."

2.4 COUNTRY FOCUS: TANZANIA

TANZANIA: QUESTION 4

"How could citizen-generated data contribute best to monitoring the Sustainable Development Goals (SDG)?"



Answers

Validating or questioning official statistics.	Identifying areas for further data collection.	Increasing data coverage.	Identifying potential programmes and policy responses.	Additional information to complement official indicators.	Providing local level context.

Opportunities.

The graph shows how we can see that *identifying areas for further data collection* and *additional information to complement official indicators* were consistently assigned high rankings. *Increasing data coverage* also received, on average, good rankings. Once again, respondents saw more potential for CGD initiatives in accommodating government needs than they did in criticising its shortcomings. "CGD should complement official statistics rather than question or validate official data", said S. Chacha (Africa Philanthropic Foundation). "Over time that may happen but that will be realistic only once CGD can correspond to official data standards."

Until that level of maturity is reached, any exposure by CGD of public shortcomings may be discredited on a methodological basis. An aggressive strategy may then harm CGD's agenda and lead to a premature end to inclusive monitoring. Globally agreed standards are thus necessary to ensure some protection against vested interests excluding citizens' voices from official statistics.

While agreement on the importance of *providing local-level context* varied, common standards would also align collection practices and, once CGD initiatives can be scaled, allow for modular datasets to or from which local data can be added or pulled. "Hyper-local collection would be a viable alternative for programming and near-real-time priority-setting and course-corrections",[13] added M. Blaser (MCC-PEPFAR).

Sectors.

"CGD can only complement sample surveys (not censuses) provided they observe principles of official statistics", argued a government representative. For S. Chacha (Africa Philanthropic Foundation) – who agreed that "CGD should not be used for census, where data is sufficient" – the sectors where CGD can have the most impact vary on a country basis. "[In Tanzania] we seriously need CGD in all areas of development monitoring, except one or two", owing to widespread data gaps.

Currently, CGD initiatives in Tanzania are tackling diverse issues, including health (UNICEF's Birth Registration System and Helpage International Tanzania),[14] human rights (Tanzania Human Rights Commission),[15] public services (Taarifa Project),[16] and land rights (Agrinfo).[17] Sauti za Wananchi, meanwhile, uses mobile phones to survey large groups of people quickly and at low cost on different topics like health, education, governance, and water and sanitation.[18]

13 For example, Premise (**http://premise.com**) combines crowdsourced data collection – users take pictures of consumer products in local stores – with machine learning algorithms to monitor consumer prices in near-real time.

14 See, respectively, **http://unicefstories.org/2016/11/30/advancing-the-birth-registration-system-in-tanzania-providing-under-five-children-their-right-to-protection/** and **http://www.helpage.org/silo/files/ocm-guidelines.pdf**

15 See **http://www.chragg.go.tz/**

16 See **https://taarifa.hackpad.com/Taarifa-Project-Overview-ZejYsiU17xQ**

17 See **http://www.agrinfo.co.tz/**

18 See **http://www.twaweza.org/go/sauti-za-wananchi-english**

Challenges.

Lack of representativity and *lack of methodological rigour* were considered the biggest challenges. The latter supports concerns that officials may dismiss unfavourable data for methodological reasons. The importance given to *lack of representativity*, on the other hand, is a reminder that, even if CGD becomes a key part of a more inclusive statistical system, failure to significantly improve the issues and populations it covers could limit its efficacy in this regard.

"We really need to make sure that we have relevant data that represents all groups within communities, but we also need to maintain flexibility for each context", stressed S. Chacha (Africa Philanthropic Foundation). He recalled how, during an event on "leaving no one behind", he was using official statistics to prove his points when a woman contested that national figures completely misrepresented the situation of her local community – regardless of how methodologically rigorous sampling may be.

Though official statistics may not capture the whole picture, CGD (especially when extremely localised) must follow standards that policymakers can realistically consider. "Rapidly conceived CGD may be informative in some cases", said M. Blaser (MCC-PEPFAR), who then added that "sufficient levels of inclusivity and rigour are a prerequisite for any convincing arguments supported by CGD."

While there was no strong consensus on other challenges, another issue emerged from the interviews: the need for improved data skills. "You may have a methodology that complies with official standards but it is useless if you do not have the skills to follow it", said S. Chacha (Africa Philanthropic Foundation.

2.5 INTERNATIONAL NON-GOVERNMENTAL ORGANISATIONS

Opportunities.

While international organisations work mainly at a macro level – in contrast with CGD's local nature – they develop national and regional indicators to monitor the SDGs and can help in designing flexible standards that span across borders.

Respondents gave higher priority to *additional information to complement official indicators* (consistently ranked first) and *increasing data coverage*. Less importance was given to *identifying potential programmes and policy responses* and *providing local-level context.*

CGD can be useful to "build consensus and catalyse new approaches to monitor policy effectiveness, and have civic initiatives take the role of government watchdogs", added one respondent. Hence it appears that international non-governmental organisations (INGOs) are less focused on the collaborative aspects of data collection and envision CGD more as a tool for independent advocacy than as one to support official monitoring.

Sectors. Interviewees mentioned several sectors, including health, education, gender, and labour. Everyone recognised that CGD should not be limited to a particular goal; rather, its potential resides in providing very localised information.

"[Sectors of activity] need to be assessed on a careful basis together with international collaborative partners", said one respondent, and "standalone efforts risk being marginalised and if not mainstreamed would appear more as data activism for data activism sake and not really a concerted effort to address the SDGs and affect policy change in an official capacity". Similarly to some respondents from pilot countries, these answers reinforce the need for consolidation of practices and strategic activism.

Challenges.

Respondents from INGOs gave the highest priority to *the nature of relationships between civil society and national statistical and monitoring professionals*, followed by (with equal ranking) *lack of representativity*, *lack of methodological rigour*, and *persistency of data coming from unofficial sources*.

3 DISCUSSION AND OPPORTUNITIES

OPPORTUNITIES AND CHALLENGES: A SUMMARY

Although CGD is at its core a tool for advocacy, currently we also need initiatives that mirror or complement official information (rather than presenting completely new narratives) to help consolidate collection and validation methods.

CGD is firstly made by and for citizens but it must also reach the attention of policymakers. It needs to cater to the needs of both audiences through a virtuous cycle based on mutual trust whereby citizens are trained to collect quality data that can be used for policymaking and, at the same time, feel that officials will consider CGD in decision-making processes.

Statisticians working for public agencies (familiar with official collection and validation methodologies) should train citizens who, in exchange, can support governmental monitoring.

CGD can have both an advocacy and an "administrative" purpose, or even both. It would then be useful to expand the current definition of CGD (in terms of what it is and what it can be used for) with "modular" blocks that can be chosen according to the type of data and its goal.

Indirect consequences of more inclusive SDG monitoring may include 1) finding better ways to communicate monitoring results to wider audiences; 2) allowing new non-governmental players into the dialogue on monitoring and implementing the SDGs; and 3) promoting data literacy and culture.

Financial resources are key to both the persistency of CGD initiatives (important to build credibility) and the quality of data they produce. Moreover, it is important that projects are replicable to prevent certain areas being ignored in data collection while others get saturated.

This section builds on DataShift's understanding of both CGD initiatives and the civic space in its pilot countries. Our aim is to build an open knowledge repository that monitoring professionals and civil society alike can use to learn from each other and initiate dialogue.

This section categorises problems as reinforced or additional barriers. The former refer to obstacles mentioned in the last section, which can be expanded on using previous research; the latter, on the other hand, represent relevant problems not covered during the interviews but can still have reflections added on how to conduct more effective advocacy and dodge potential bumps. The box below summarises the most important insights that emerged in the previous section.

3.1 ARGENTINA

REINFORCED ISSUES.
FINANCIAL STABILITY // OPEN TOOLS AND STANDARDS // SHARING KNOW-HOW.

Financial scarcity is a core issue for Argentinian organisations. A study by CIVICUS shows how the biggest threat to the stability of more than 100,000 CSOs in the country relates to insufficient funds and resources.[19] Restrictions to international funding further complicate matters, especially for organisations with little experience in terms of the financial requirements imposed by the government.

New approaches (like crowdfunding) can help but initiatives should prioritise understanding the financial system, perhaps through collaboration with other CSOs. Open source technologies and volunteers may also help in relation to containing costs. Underestimating financial stability, however, is likely to affect both data collection and community-building efforts. This will affect the credibility and the impact of CGD, undermining the efforts of those who voluntarily participated.

Open source software also makes initiatives more replicable. When pushing for common standards with governmental actors, civil society stakeholders should stress the importance of openness. Accessible tools – software, formats, and licences – may not be enough to replicate success stories about the countries: when possible, resources should also be devoted to documenting how challenges were addressed and what lessons were learnt.

CGD initiatives are often pioneers and this can isolate them from other actors in their field. While their strength lies in the creative leveraging of technology that can also help non-data-driven initiatives, CGD projects may also need support from other CSOs.

Think, for example, about political sensitivity: if CGD is not properly curated before it gets published it may harm collaborators and their communities. Unexpected backlashes (e.g. as a result of failing to comply with privacy laws) can damage those who were supposed to receive help in the first place.

19 Balian et al. (2011).

Seeking expertise from other organisations (such as those experienced in human rights advocacy) can help prevent negative consequences both for the initiative and for its users.[20]

Hence efforts should be devoted to promoting communication, not only between civil society and public institutions but also within civil society itself. Forums can be a chance for CSOs to align their agendas and learn from each other about sensitive issues. This would also provide a support infrastructure to mobilise resources if things go wrong.

ADDITIONAL ISSUES . MEASURE IMPACT // UNITE STAKEHOLDERS.

Previous research shows that many CGD initiatives in Argentina do not prioritise systems to track their user base.[21] Though recording how people participate in data collection and their characteristics may conflict with principles of openness, when done in a transparent way (e.g. stating which data is collected and why, providing opt-out options, etc.) it can help address data representativity and validation without adding substantial burdens – since profiling mechanisms and checks can be embedded in both apps' web pages.

> *"Though hard to measure, it appears that the sense of empowerment comes more from the act of participating (and through learning from practice and observation of others) than from data production."*[22]

CGD initiatives have higher chances of making citizens' voices heard when they are supported by a strong community. Meanwhile, users from established initiatives reported a sense of empowerment from participation as they identified themselves as part of a wider collective. Initiatives should however, be aware that their government may not react to their advocacy, or that it will do so in negative ways. It could, for example, label initiatives that criticise its programmes as political opponents. Isolated organisations will not have the means to face this resistance, but engaged users and collaborative organisations can count on stronger political standing.

20 While CSOs in Argentina can collect and use data without particular restrictions there are laws that regulate the collection of sensitive data (e.g. racial and ethnic origins, political opinions, and sexual behaviours). Legal expertise is useful to make sure regulations are fully respected. For further information see Kobus and Zeballos (2015) and Christie et al. (2013).

21 Fressoli et al. (2016).

22 Ibid.

3.2 KENYA

REINFORCED ISSUES.
IDENTIFY TARGET AUDIENCE AND BUILD TRUST // BALANCE SIMPLICITY AND ACCURACY.

Previous research confirms how lack of trust between government and civil society actors is undermining the impact of CGD.[23] Negative or suspicious attitudes can call into question the validity of unofficial data. Whether these owe to genuine methodological concerns or a fear that CGD may harm the reputation of government functionaries, it can lead to officials discouraging further activity.[24] On the other hand, trust building can also help in identifying the right offices and individuals to communicate with.

When CGD initiatives focus only on delivering data to policymakers, however, they will prepare it in ways citizens may not understand. Previous reports show that some CGD initiatives were considered government-led projects and people expected them to deliver on promises that did not come under their responsibility (especially relevant for projects measuring perceptions about public services).

Lack of methodological rigour, especially when sampling, is another factor reducing CGD attractiveness for official statisticians. Most of the initiatives in previous reports we examined lack rigorous collection methodologies.[25] This, however, may prove advantageous in local contexts where complex methods are not implementable. The "one-size fits all" solution is not, as previously said, the best option for data collection in areas with varying conditions (such as related to internet connectivity and literacy levels), trained statisticians and civil society mediators should account for this when setting common standards.

ADDITIONAL ISSUES.
ACHIEVING FINANCIAL SUSTAINABILITY.

Most Kenyan CGD initiatives we examined rely on donor funding, whose limited duration often affects their potential for improvement. Projects may seek to improve their collection, processing, and dissemination skills but these needs may clash with donors' priorities, such as for a return on investment. Though most initiatives receive funding from international resources,[26] many lack skills to make effective grant applications.

23 Oduor-Noah et al. (2016).
24 See the Centre for Advocacy and Research Development (CARD) case study in ibid.
25 Ibid.
26 Ninety-one percent of all NGOs in Kenya are funded by international sources, eight percent are funded by private local sources and only one percent receive their funding from the government.

3.3 NEPAL

REINFORCED ISSUES.
UNCLEAR CIVIC-OFFICIAL RELATIONS // STRICT REGULATIONS.

Interviewees conveyed, more or less clearly, a common message: CGD initiatives in Nepal should aim at preparing favourable foundations on which further projects can be built. CSOs should not underestimate the importance of good relationships with officials both at the local and at the national level; this is crucial to gain know-how about data collection, as well as to avoid potential conflicts with public officials.

Previous work on civic space in Nepal reports how "there is much room for improvement concerning the relations between the government and CSOs"[27] and currently the situation is characterised by "little trust and constructive engagement."[28] In terms of data collection and use, the latest Statistical Act reserves the right to coordinate professional data collection to the Central Bureau of Statistics (CBS). Anyone who intends to collect, use, or publish data must first obtain permission from the CBS.

As regards financial sustainability, CSOs in Nepal must deal with strict barriers that regulate resource inflows and impose unfavourable tax incentives. The difficulty involved in receiving money, materials, and even volunteers from abroad limits the scope of activity, as well as the chances for partnership with international CSOs.[29] This is especially relevant as few initiatives in Nepal, both those of civil society and those of government, have adopted the SDG framework.[30]

ADDITIONAL ISSUES.
RESOURCE-DEMANDING PRACTICES // UNCLEAR MANDATES.

As in other countries, Nepalese CGD initiatives are pioneering data culture in their communities. Previous research also shows that low literacy levels make data collection more labour-intensive in rural areas, mistakes are more frequent when volunteers have to fill forms for respondents, and lack of established verification methods further decreases perceived data quality.[31] Moreover, current methods (like data verification through phone calls) have often proven hard if not impossible to carry out.

Some projects have reported that local communities, especially in rural areas, have mistakenly interpreted their reporting as part of official development projects and, consequently, expected data collection to be followed by service provision. While official endorsement would legitimise CGD-related operations, it should also clarify how initiatives led by civil society are not responsible for governmental mandates.

27 Buentjen (2014).
28 Bhargava (2015).
29 ICNL (n.d.).
30 Budhathoki et al. (2016).
31 Ibid.

3.4 TANZANIA

REINFORCED ISSUES.
BUILD CREDIBILITY // UNCERTAIN CIVIC SPACE.

Interviews asserted the importance of embracing a cooperative rather than competitive attitude towards institutional actors in Tanzania. The "experimental feel" attributed to CGD initiatives can undermine their actions. As previously said, if institutional actors decide to shut their ears to citizens' voices efforts will be pointless.

Credibility must then be achieved in different ways. First, a clearly defined theory of change can provide a long-term vision for each initiative. This can help manage expectations too: previous reports show that local communities may misinterpret the mandate of CGD initiatives and expect them to deliver public services rather than to support their claims.

Second, the space for civic action in Tanzania is uncertain. "[The] Statistical Act allows the National Bureau of Statistics to work with any CSO or any other organisation to monitor not only SDGs but also any other national and international development frameworks", said one government official we interviewed. The same bill, however, restricts the publication of data to government and limits the operations of CSOs and academic researchers.

Moreover, the government has, through the NGO Board, the power to dissolve CSOs.[32] Therefore while nurturing relations with government is key to CGD survival, and limiting these projects to less-controversial areas can increase the likelihood of the data they produce being used by officials, this is of course immensely problematic from both an accountability and issue-coverage perspective.

ADDITIONAL ISSUES.
BALANCE DATA PACKAGING // INTERNAL EVALUATION // DIVERSIFY SERVICE PORTFOLIO.

CGD initiatives should strive to balance rigour and accessibility when presenting their data. On the one hand, CGD that targets policymakers may not always be understood by the same people that helped collect it, thus undermining their sense of purpose. On the other, many ongoing projects focus on measuring perceptions of public services, leaving data subject to criticism on the basis of low methodological rigour and negative attitudes from politicians who see their authority questioned.

As regards sustainability, organisations lack internal evaluation mechanisms. Twenty percent of Tanzanian NGOs have reportedly suffered from governmental attempts to suspend them on the basis of unfulfilled reporting requirements.[33]

32 TANGO (2013).
33 Ibid.

Moreover, they also often suffer from a lack of funding. There are few examples of organisations that have been able to diversify their products and services and charge for data-driven consulting. This, however, will not fit all projects, and many will continue relying on donor funding (which often lasts for only one year). In addition to undermining their stability, limited funding blocks resources to improve internal skills, do lobby work, and engage stakeholders. Furthermore, several initiatives struggle to make effective grant applications.

4 IMPLICATIONS FOR NATIONAL MONITORING STRATEGIES

In addition to surveying current perspectives about opportunities and challenges for CGD, our interviews also explored the interest and readiness of government actors in terms of integrating CGD within their processes. More specifically, we asked participants to rank actions according to their potential to achieve short-term impact. In the next pages we integrate these results with the findings from previous sections to extract context-relevant strategical implications.

4.1 ARGENTINA

The biggest threat to Argentinian CSOs relates to financial sustainability. Though cases exist of government critics being branded political opponents (resulting, for example, in journalists losing their jobs[34]), finding sufficient and long-term resources remains the biggest challenge for new CSOs.[35]

Among the countries examined, Argentina seems to be the one with the highest potential for CGD to promote an alternative to official narratives. However, this does not mean CGD initiatives in Argentina should expect to make a substantial impact without cooperating with the government to set commonly endorsed methodologies that include SDG metrics.

Endorsing a participative SDG agenda.

There is a sense among civil society actors that the government is still working on adapting its development agenda to the SDG framework. For example, O. Minatta's organisation (ACIJ) tried to request information from official actors about the new metrics being designed and did not receive a response. This was further exacerbated at a recent High Level Political Forum (HLPF), government did not disclose much information about its current work on new SDG-related indicators.

34 CPJ (2015).
35 Fressoli et al. (2016).

"At the global level there are some resolutions from the UN that express the necessity of civil society engagement. However, this has not been converted into a clear agenda that ensures citizen participation", O. Minatta added. While lack of consensus may justify the reluctance of officials to share their work, it also represents an opportunity to collect insights from civic actors who understand the global agenda as well as Argentina's local context and the issues facing the communities they represent.[36]

Sharing methodological know-how.

CGD can have impacts on a wide range of governmental activities, though it must first be recognised and legitimised by public figures. Methodological rigour and data accuracy are crucial, especially in vulnerable communities where proper data collection is lacking. R. Borrmann (Ministerio de Modernización y Gobierno Abierto) believes priority should be given to promoting data literacy among groups of citizens rather than CSOs.

Authorities should educate citizens directly about official data collection methodologies. CSOs, on the other hand, can serve as mediators, since they are usually in a better position to engage and understand underrepresented groups.

More collaboration is also needed between CSOs to leverage complementary strengths and avoid duplication of efforts. For instance, N. Aquilino (CIPPEC) thought that, even if an organisation does not have data experts, it can "outsource" the necessary skills from partners within civil society or from the private sector (provided sufficient funds are available).

Collaboration should be mutual: initiatives that possess data-related skills should also seek guidance in navigating the legal landscape and institutional operations. Consequently, data-driven advocacy could be packaged according to institutional working principles to increase their impact and avoid unlawful data collection that can harm an initiative's long-term sustainability.

4.2 KENYA

The state of digital rights in Kenya does not look particularly promising. The country's Data Protection Bill, drafted in 2009, is an all-encompassing law that should protect citizens' privacy but presents potentially harmful aspects towards data rights.[37] The African Union (AU) Convention on Cyber Security and Personal Data Protection, on the other hand, is supposed to protect human and digital rights

36 For example, she noted that Argentina had recently presented its environmental reporting plans for the year: "It would be interesting to analyse what kind of information countries voluntarily report compared with the different SDGs that are planned to be revised in the next HLPF."

37 Article 19 (2011).

across its member states. But this will not come into force until all 15 AU members implement it. This leaves each government with the right to pass legislation that may have undesirable effects both on individual freedoms and on civic space.

Together for flexible standards.

Working towards the development of standards scored the highest in our aggregate results. This requires consistent and well-intentioned collaboration between government and civil society actors, which in turn can help identify the best venues to raise CGD impact within government and foster trust. Mutually recognised standards will not be easy to achieve, however, especially considering the need for flexibility to adapt into areas where data collection will be more difficult.

Multi-stakeholder workshops to jointly develop data collection and use guidelines also scored high. Consistent discussion and trials will likely be necessary and government may show aversion towards risk of failures. Hence it is important that CGD is presented as a way to reduce administrative and monitoring burdens rather than strengthening civil advocacy. "We need to define CGD better: we are not only collecting data to demand accountability, benefits for government must be clearly stated", said K. Rono (Development Initiatives).

Both *secondments and fellowships* and *collaborative monitoring* have lower aggregate scores, suggesting these actions should be undertaken only after solid foundations are set. Finally, no clear answers emerged about the integration between CGD and official data - that is, no option consistently ranked as high or low priority. Terms and modes of collaboration between government and civil society actors must be consolidated before data integration can be discussed.

4.3 NEPAL

The civic space in Nepal remains uncertain as it is unclear how many CSOs operate in the country (estimates range from 6,000 to 84,000) and the government is still defining its relationship with these organisations.[38] Though the representatives we interviewed expressed much enthusiasm for the perspective of including civic actors in data collection to monitor the SDGs, we cannot ignore that many Nepali CSOs reportedly faced a hostile environment.[39]

Undefined relations, lack of experience of constructive negotiations, and restrictions on funding (see previous section) mean CGD initiatives in Nepal need to initiate a mutually beneficial dialogue with their government. This may have to [redacted]mited to uncontroversial issues, however.

38 ICNL (n.d.).
39 Ibid.

Planting the seed.

When asked about the actions that should hold higher priority in terms of promoting a pluralistic agenda for sustainable monitoring, Nepali respondents gave *data catalogues* the highest aggregated ranking, confirming how Nepalese civil society and institutions should join forces to identify data gaps and ways to move forward together.

Interviewees focused more on preparing the grounds for collaboration rather than engaging in data collection: *common standards*, *multi-stakeholder workshops*, and *capacity building programmes* gained higher priority than *collaborative monitoring* and *secondments and fellowships*.

Offer help, demand training.

D. S. Shrestha (NGO Federation of Nepal) noted how SDG indicators would require "lots of data, and disaggregated data too". This will prove challenging as Nepal's CBS "does not have the necessary capacity to sustain data collection for SDG metrics". However, it was not possible to get a CBS representative to respond to this claim.

The need for dialogue to find common ground was confirmed when K. H. Baskota (National Information Commission) admitted that, while being "very interested in working with CGD", he had a "lack of knowledge" on the topic.

Two opportunities therefore emerge. The first relates to CSOs offering their help to governmental agencies to collect data for the SDGs. While implying an initial compromise – data collection would be limited to non-conflictual issues – this would promote the development of standards, know-how transfer, and positive relationships. The second opportunity, which comes as a consequence of the first, is about creating more engagement between citizens, CSOs, and local officials.

While official monitors may welcome the offer to help in SDG monitoring activities, it is not granted that they will take it seriously. As previously said, governmental practices in data collection demand resources that few organisations or citizen groups have. Technology can help (when available) but official practices still need to be adapted to practical constraints. Less detailed yet valid methodologies are necessary and, while CSOs' community reach can help mediate relations, citizens should receive training directly from monitoring officials.

4.4 TANZANIA

The 2002 NGO Act strictly regulates where CSOs in Tanzania can operate and which activities they can engage in.[40] Moreover, as previously noted, the NGO Board has the power to dissolve organisations, while the Statistical Act, the Media Services Bill, and the Cybercrime Bill risk undermining both freedom of expression and privacy.[41]

Given this fragile civic space and the relative novelty of CGD, initiatives that openly criticise the government will likely be reprimanded in ways which hinder them from achieving impact. It is thus necessary to create foundations (established methodologies, collaborative channels of communications, and solid partnerships) so that CGD can grow and establish itself as an universally recognised viable source of information for SDG monitoring.

While *standards* were consistently ranked as high priority (between second and third position) there is less consensus on other issues. Though this makes it harder to identify actions that should be undertaken first, it does not mean that participants consider issues irrelevant. As a matter of fact *capacity building programmes* ranked between second and fourth position whereas *collaborative data catalogs* came in slightly lower, at between third and fifth.

Consolidate, consolidate, consolidate.

S. Chacha (Africa Philanthropic Foundation) put a great deal of importance on documenting practices and building networks. "[While] we are not collecting data ourselves we are consolidating different ways of collecting CGD to promote standardisation at the national level", adding that "there is no point in having many initiatives if these do not speak to each other."

Much of his organisation's work is about consolidating practices. Making civil society actors talk to each other will help them build a common front and replicate the methods that prove more effective. "Standardisation is really key. If I am able to collect data on women land ownership rights in a Tanzanian village, that data should really be a building block in the national overview on the issue. I need to be able to have similar blocks that I can use to build a national overview on the problem."

Action, and measuring it.

Official endorsement and recognition ranked consistently high in {{graph 8b}}. While common standards can be mutually agreed, endorsement and recognition require supportive political figures. Whether the institutional environment understands the

40 Kepa (2013).
41 TANGO (2013).

potential benefits of CGD, assessing its impact is key to ensure its acceptance. In this regard, one respondent reminded us that "as much as we think of CGD as a new concept, data is not a new concept! What's missing is how can we push CGD usage to drive the impact we'd like to see."

4.5 INTERNATIONAL NON-GOVERNMENTAL ORGANISATIONS

More cooperation through successful pilots.

As regards actions to promote pluralistic monitoring, all respondents from INGOs agreed that national and international-level efforts can help consolidate civic–public partnerships and scale data collection. As a first step, however, pilots are necessary to offer models on which INGOs can build their advocacy work.

According to one respondent, "before we talk about data collection, there has to be a core agreement on methodologies. The reconciliation of sampling issues and transparency of sharing frames; ease of access and freedom of access (and confidentiality) need to be discussed." Moreover, while strategies will vary in each country, generally CGD initiatives should focus on providing solutions to ease the work of official monitors. In the eyes of institutional players, if inclusive collection implies additional burdens, both financial and political commitments can be easily undermined.

5 CONCLUSIONS

In our quest to understand the perspectives of statistics and monitoring professionals on CGD we identified a number of overarching issues present in most pilot countries. The relative novelty and pluralistic nature of CGD mean public officials can dismiss any that questions their authority. Their criticism may stem from genuine concern about collection methodologies, accuracy, privacy risks, and the compatibility of CGD with established (and often inflexible) official procedures. Additionally, though, CGD may also highlight shortcomings in public service provision or vested interests, giving rise to a less legitimate (but nonetheless effective) resistance.

Non-governmental stakeholders may have to accept their role as newcomers in development monitoring, otherwise the efforts of citizens who voluntarily contribute to CGD production could have little impact – and thus an opportunity to raise citizen voices and influence development policy will be missed. Civil society in some cases can present itself as supportive of the government rather than as a source of dissent.

Governments in our pilot locations are affected by significant data gaps that will likely prevent proper monitoring of the SDGs. By offering to help fill these gaps, civil society can establish mutually beneficial partnerships that will elevate CGD to being a recognised source of information.

Yet there appears to be enthusiasm within government circles for engaging with CGD, which presents civil society with an important opportunity, especially for less well-established projects. Therefore, while the road is long and contains many barriers, the following approaches, if promoted and applied by civil society actors (and cooperative governments), can lead to the constructive dialogue and a collaborative approach required to improve the effectiveness and sustainability of CGD projects.

STEP 1

ESTABLISH GOOD RELATIONS WITH LOCAL AND NATIONAL GOVERNMENTS TO PROMOTE MUTUAL TRUST

The relationship with statistics and monitoring professionals is crucial as it will affect which data should be collected, what methodology to follow, and what standards and licences to adopt.

Jointly identify data gaps that are relevant to SDG monitoring and where CGD initiatives can help institutions.

- Agree on a common framework for data collection and processing:
  - How can official methodologies be adapted to contexts where resources are scarce while maintaining an acceptable level of data accuracy?
  - What data formats and licences should be used? Open source tools and licences to contain costs and facilitate replicability and complementarity should be promoted.
- Agree on metrics to evaluate the impact of each initiative (e.g. user-based, communities involved, spin-off projects, etc.).
- Request/organise training and (when relevant) access to new technologies
  - CSOs can help mediate as they have a better grasp of the local context and contacts with the community. They may also be better at presenting technical material prepared by official statisticians to untrained audiences.
- CSOs should look for ways to present your proposal as something that would ease the government's work rather than adding a burden!

STEP 2

ESTABLISH SOLID PARTNERSHIPS WITH OTHER CIVIL SOCIETY ACTORS

As pioneers of data-driven policy – which is still misunderstood even in developed economies – CGD initiatives may feel isolated from other CSOs. Data is not a new concept, however, and the civic space is populated by many actors that address complementary issues.

- Find partners that can value your innovative edge and assist with the skills you lack, such as:
  - Legal issues;
  - Financial reporting;
  - National and international grant applications;
  - Sensitive issues (can the data I am collecting end up harming the communities I intend to help?)
- Experimentation will lead to unexpected turns. Build a network for mutual support and resource mobilisation when things go bad. Especially if civic space is uncertain.
- Nurture your community. Your organisation is not about data; data is a means to solve the issue that affects the community you have engaged. This affected community is thus united by a common problem.

- Foster a sense of community and belonging because it will empower people by making them feel part of a bigger collective.
- Beware of unfounded expectations: CGD initiatives that monitor public services do not have the mandate to provide those services.
- When possible prioritise underrepresented communities.
- Promote a broader data culture that goes beyond issues you are tackling (e.g. disseminate information about digital footprints).

STEP 3

SEEK INTERNATIONAL SUPPORT

- INGOs or international networks/platforms can help streamline agreement on standards and methodologies:
  - In the countries examined most government agencies have not yet adopted the SDG framework. International actors may have a better understanding of the metrics and can help you adapt them to your local context.
  - They are usually well versed in legal rules and procedures, so if you comply with their standards (e.g. in privacy) you can probably also promote them in your country.
  - They can help you understand in which areas to focus and how to find support from other actors in your field so your initiative does not get marginalised.
- These actors can also help to fund projects and/or provide guidance to obtain international funding.

STEP 4

OPEN YOU WORK!

- Replicability is crucial. CGD's impact will significantly increase when different communities collect the same data in comparable ways.
- Using open formats, software, and licences also nudges others (both government and civil society) to do so.
- Consolidate practices. Opening your code and data gives up the ingredients to the rest of the community but says little about the steps you followed. When possible, you should devote resources to documenting how you set up your initiative, the challenges you faced, how you addressed them, and what lessons you learnt from your mistakes. Storytelling can give important insights to those wishing to emulate you.

6 RECOMMENDATIONS FOR DATASHIFT AND PARTNERS

EXPAND THE CGD DEFINITION AND MAKE IT "MODULAR" SO IT CAN BE ADAPTED TO DIFFERENT NEEDS

Currently CGD is defined as "*data that people or their organisations produce to directly monitor, demand, or drive change on issues that affect them.*" This is a new concept, often unknown in official circles, and thus there is an opportunity to present CGD in a way that appeals directly to the listener.

- When a CGD team approaches its local government it can highlight, for instance, how CGD can help in local monitoring processes, easing the work of public officials. "*CGD is data that people or their organisations produce to fill data gaps and help local authorities better understand the current situation on social, economic, and development matters that are relevant to the SDGs.*"
- When approaching other CSOs or citizen communities CGD could be presented as a way to make their claims more legitimate. "*CGD is data that you can produce to strengthen your advocacy with evidence whose validity is comparable to the information government officials use to draft policy.*"

EXPAND THE LEARNING ZONE TO INCLUDE A "TOOLBOX" DESCRIBING THE DIFFERENT TECHNOLOGIES AND METHODS TO LAUNCH A NEW CGD PROJECT

The toolbox should cater to different levels of expertise: expert users should be redirected to more in-depth sources, whereas those with less technical skills should be able to access simplified material. It could build on already existing work, such as:

- Citizen Science initiatives on data collection methods: **Mobile Data Collection Guide** by School of Data.
- Open Data community on how to publish data in an reusable and interoperable way: Data on the Web Best Practices by W3C.
- The Network of Innovators by GovLab at NYU where users can post a request for help in a specific field of open government practices and their needs should be matched to the experts that are part of the community.
- Guidance to apply for grants and case studies on sustainable business models.

DataShift's events and Community calls are crucial to keep building a sense of community and foster knowledge exchange. Perhaps DataShift should therefore explore moving from a mailing list to a forum-like platform.

DRAFT COMMON STANDARDS

It will likely prove hard for individual initiatives to convince government officials to devote time to set commonly agreed standards, especially when it comes to balancing accuracy and flexibility.

The Global Partnership for Sustainable Development Data (GPSDD), however, has the leverage to start the discussion with its member organisations (especially NSOs). Furthermore, when actually engaging with government actors on this agenda, a concise "introductory" document that CGD initiatives and others can bring to the table when offering their help to local officials is a crucial tool for starting the conversation on the right footing.

REFERENCES AND FURTHER READING

Article 19 (2011). "Kenya: Draft Data Protection Bill critically limited". 7 November. Available at **https://www.article19.org/resources.php/resource/2825/en/kenya:-draft-data-protection-bill-critically-limited**

Balian, B., Cecconi, E., & Cao, C. (2011). "Civil society in Argentina at the bicentennial". CIVICUS Civil Society Index for Argentina (2008-2011). Available at **http://civicus.org/images/stories/csi/csi_phase2/argentina%20acr.pdf**

Bhargava, V. (2015). "Constructive engagement processes between governments and CSOs in Asia: Are we getting results? Open Government Partnership Blog, 4 February. Available at **http://www.opengovpartnership.org/blog/vinay-bhargava/2015/02/04/constructive-engagement-processes-between-governments-and-csos-asia**

Budhathoki, P., Wiele, Q., & Johnson, C. (2016). "The impact of citizen-generated data nitiatives in Nepal." DataShift. Available at **http://civicus.org/thedatashift/wp-content/uploads/2016/05/CGD-impact-report-Argentina.pdf**

Buentjen, C. (2014). "Engagement between governments and CSOs: Are we getting results?" ADB Blog, 14 October. Available at **https://blogs.adb.org/blog/engagement**-between-governments-and-csos-are-we-getting-results

Christie, A., Craig, C., Halpert, J., Jansen, T., Kashatus, J. M., Lucente, K., ... Van Eecke, P. (2013). "Argentina: Data protection laws". *World handbook: Second edition*. Sydney: DLA Piper Australia. Available at **http://www.mondaq.com/x/230846/data+protection/**Data+Protection+Laws+of+the+World+Handbook+Second+Edition+Argentina

Cornforth, J., Adieno, D., & Moraes, C. (2015). "Making use of citizen-generated data". DataShift. Available at **http://www.data4sdgs.org/s/Data4SDGs_Toolbox-Citizen_Generated_Data_for_SDGs.pdf**

CPJ (2015). "Journalist flees Argentina after reporting on prosecutor's death". Available at: **https://cpj.org/2015/01/journalist-flees-argentina-after-reporting-on-pros.php**

Fehling, M., Nelson, B. D., & Venkatapuram, S. (2013). "Limitations of the Millennium Development Goals: A literature review". *Global Public Health*, 8(10), 1109–1122. **https://doi.org/10.1080/17441692.2013.845676**

Fressoli, M., Arza, V., & del Castillo, M. (2016). "The impact of citizen-generated data initiatives in Argentina". DataShift. Available at **http://civicus.org/thedatashift/wp-content/uploads/2016/05/CGD-impact-report-Argentina.pdf**

Fukuda-Parr, S. (2013). "MDGs: Facing up to the limitations of global goal setting". The Guardian, 20 May. Available at **https://www.theguardian.com/global-development-professionals-network/2013/may/20/millennium**-development-goals-targets-global-development

Gray, J., Lämmerhirt, D., & Bounegru, L. (2016). "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**

ICNL (n.d.). "Civic freedom monitor: Nepal". Retrieved 7 December 2016. Available at **http://www.icnl.org/research/monitor/nepal.html**

Kepa (2013). "Reflections on the state of civil society in Tanzania". Available at: **https://www.kepa.fi/tiedostot/cs_tanzania_2013.pdf**

Kobus III, T. J., & Zeballos, G. S. (2015). *2015 international compendium of data privacy Laws*. Cleveland, OH: BakerHostetler. Available at **http://towerwall.com/wp-content/uploads/2016/02/International-Compendium-of-Data-Privacy-Laws.pdf**

Landry, J.-N., Webster, K., Wylie, B., & Robinson, P. (n.d.). "How can we improve urban resilience with open data?" Open Data for Development. Available at **https://drive.google.com/file/d/0B8BZxtR6WDBJcHYzallaZFNBRFU/view**

Morales, L. G., Hsu, Y.-C., Poole, J., Rae, B., & Rutherford, I. (2014). *A world that counts*. UN Independent Expert Advisory Group. Available at **http://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf**

Oduor-Noah, L., Maina, E., & Simeoni, C. (2016). "The impact of citizen-generated data initiatives in East-Africa". DataShift. Available at **http://civicus.org/thedatashift/wp-content/uploads/2016/04/CGD-impact-report-EastAfrica.pdf**

Southasia.com.au (2015). "Nepal ranked one of the worst countries for philanthropic freedom". 20 June. Available at **http://www.southasia.com.au/2015/06/20/nepal-ranked-one-of-the-worst-countries-for-philanthropic-freedom/**

TANGO (2013). "Reviewing Tanzania's non profit legislative regime and need for a new regulatory framework". Available at **http://eacsof.net/upload/REPORTS/Analysis%20Civic%20Space/TzNGORegimeReport-FinalSept2013.pdf**

UN (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. Available at **https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf**

Visit Nepal (n.d.). "Nepal non-governmental organizations". Retrieved 7 December 2016. Available at **http://www.visitnepal.com/nepal_information/ngo_in_nepal.php**

ANNEX 1

INTERVIEW QUESTIONS

Citizen-generated Data (CGD): There has recently been a proliferation of citizen-generated data produced by NGOs and citizen groups through crowdsourcing mechanisms, citizen reporting initiatives, informal web surveys and consultations, or mobile phone interactions.

This data produced great enthusiasm for its potential to raise citizens' voice and to contribute to the "data revolution", but can also be criticised for its lack of representivity or statistical rigor.
For more on CGD, see **http://civicus.org/images/ER%20cgd_brief.pdf**

- **1. Are you familiar with the term citizen-generated data as discussed above? Have you come across other ways to define it?**
- **2. Have you ever worked directly with CGD data sets, or with civil society organisations (CSOs) that produce them? If not, are you interested in doing so?**
- **3. Are there official (e.g. written statements) or unofficial plans to employ CGD within your organisation for SDG monitoring?**
- **4. How could citizen-generated data contribute best to monitoring the Sustainable Development Goals (SDG)?**
  (Please rank in order of priority and leave blank if irrelevant. The same rank can be assigned to two actions.)
  - a. validating or questioning official statistics
  - b. identifying areas for further data collection, including disaggregated data
  - c. increasing data coverage in contexts where limited resources or other considerations limit available data
  - d. identifying potential programme and policy responses to SDG implementation challenges
  - e. casting light on additional information to complement official indicators
  - f. providing local-level context

- **5. What are the main obstacles to using CGD in SDG monitoring? How serious do you think the below challenges are?**
    a. lack of representivity
    b. lack of methodological rigour
    c. comparability and coverage
    d. nature of relationships between civil society and national statistical and monitoring professionals
    e. are there other important obstacles?

- **6. a) Is there any particular domain(s) where you think CGD could complement official data (e.g. census, health, education, etc.)? b) Is your organisation already collecting and using CGD for those (or other) topics?**

- **7. Are you working with civil society organisations to help you implement and/or monitor the SDGs? If so, how?**

- **8. What do you consider the most important approaches that should be explored in order to further a pluralistic approach to SDG monitoring that promotes citizens' voices and perspectives?**
  (Please rank in order of priority and leave blank if irrelevant. The same rank can be assigned to two actions.)
    a. collaboration between official and civil-society data producers
    b. Integrating CGD into government datasets

- Additionally, document for the compilation:
    - Name:
    - Email:
    - Scope of work (multiple choice): local, state or national
    - Agency/ministry:

Federico Piovesan,
Statistical perspectives on citizen-generated data.

For more information, visit **www.thedatashift.org**
or contact **datashift@civicus.org**
First published, March 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/

Printed on recycled paper/FSC.
Graphic design by Federico Pinci

DataShift seeks to foster and inform dialogue between civil society and governments about the practical ways citizen-generated data (CGD) can contribute to action on sustainable development, including the implementation and 'follow-up and review' of the global Sustainable Development Goals. In this report we delve into the opportunities and challenges involved in using CGD to support government-led efforts to catalyse and monitor progress on sustainable development.

We conducted a series of interviews to investigate how public data producers and users (such as National Statistics Offices and policymakers) perceive CGD. Participants included statisticians, public officials, and other relevant stakeholders working at the local, state, national, and global levels. While focusing on DataShift's pilot locations (Argentina, Kenya, Nepal, and Tanzania), our study also brought in contributions of experts from international organisations – such as Open Data Watch, Paris21, and the UK Department for International Development.

Participants were asked to provide their perspectives on whether CGD can become a reliable complement to official data and help inform policy decisions, where and how governments can benefit most from CGD, the challenges to accepting CGD within institutions and the contingency actions government and civil society should focus on in the short term.

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**

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.

CIVICUS | Open Institute | THE ENGINE ROOM | Wingu Tecnología sin fines de lucro