



STATISTICAL PERSPECTIVES

ON CITIZEN GENERATED DATA

Federico Piovesan



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.

BACKGROUND

DataShift seeks to foster and inform dialogue between civil society and governments about the practical ways citizen-generated data (CGD)¹ can contribute to action on sustainable development, including the implementation and ‘follow-up and review’ of the global Sustainable Development Goals. This report delves into the opportunities and challenges involved in using CGD to support government-led efforts to catalyse and monitor progress on sustainable development.

A series of interviews were conducted 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.

METHODOLOGY AND RESULTS

In total 17 interviews were conducted with a mixed group of participants, including those from both within and outside government, as well those familiar with CGD and those less so.

To begin with, a number of barriers to the integration of official data and CGD were identified, based on the experiences of DataShift, as supporters and promoters of CGD initiatives (summarised as Challenges in the box below). Participants were then surveyed to reveal how they perceived CGD, the opportunities it presents us with, as well as challenges around its use. More specifically, interviewees’ opinions were sought on 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.

1 Citizen-generated data 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.

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 significant 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 catalogues 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.

Furthermore, a number of indirect consequences of more inclusive SDG monitoring that utilises CGD were also identified:

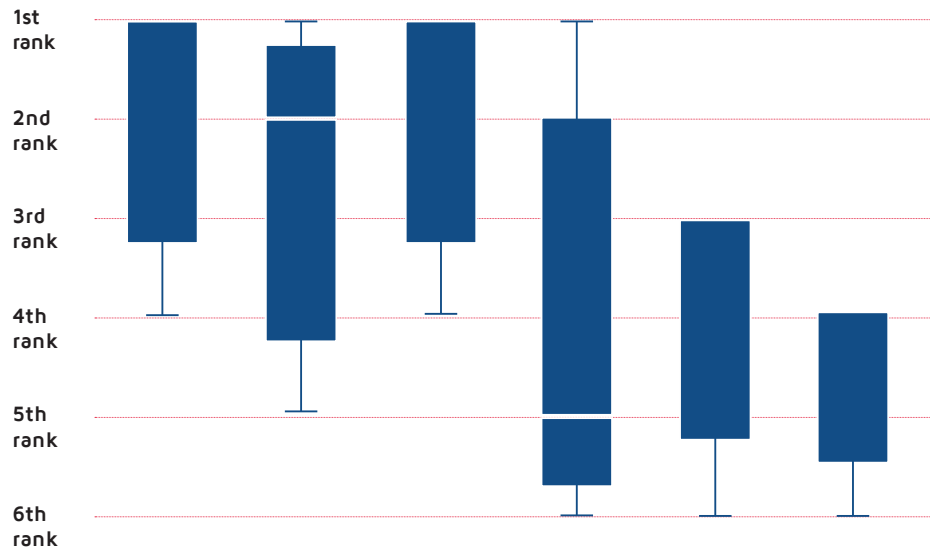
1. Finding better ways to communicate monitoring results to wider audiences
2. Allowing new non-governmental players into the dialogue on monitoring and implementing the SDGs
3. Promoting data literacy and culture. Financial resources are key to both the persistency of CGD initiatives (important to build 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.

COUNTRY FOCUS:

“How could citizen-generated data contribute best to monitoring the Sustainable Development Goals (SDGs)?”

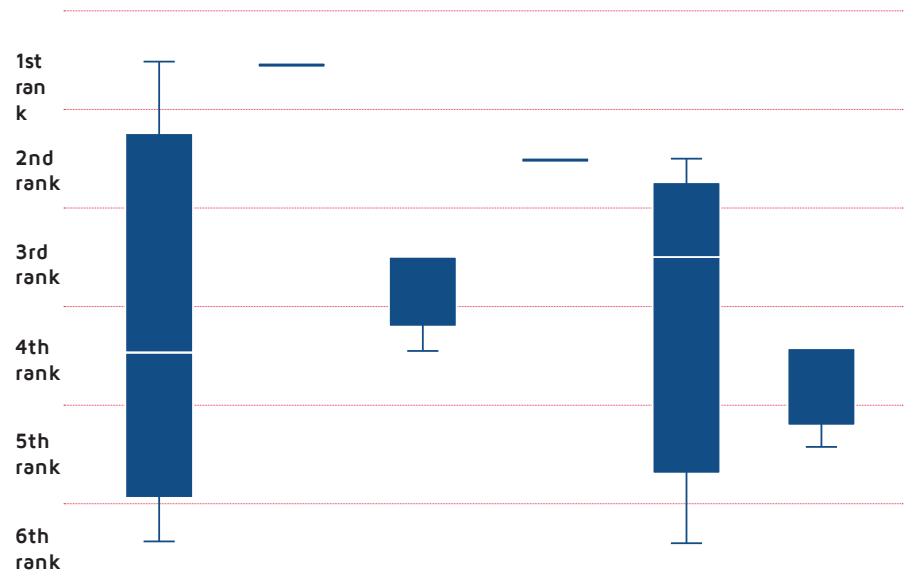
Kenya



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 |
|---|---|--------------------------|---|--|-------------------------------|

Nepal



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 indicators | Providing local level context |
|---|---|--------------------------|---|---|-------------------------------|

CONCLUSION

The interview process revealed a number of overarching issues present in most DataShift pilot countries.

The relative novelty and pluralistic nature of CGD means it can be easily dismissed by public officials can dismiss it if they feel it challenges their authority, suggesting that **greater understanding and trust around the utility of CGD is needed**. This may stem from genuine concerns 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 equally, if not more challenging) form resistance.

Governments in our pilot locations are affected by significant data gaps that will inevitably hinder effective 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. Although CGD is a key tool for advocacy, this can involve complementing and verifying official information (rather than presenting completely new narratives), which can in turn help to consolidate collection and validation methods.

Therefore, while CGD is firstly made by and for citizens, **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.

Finally, **governments appear to be enthusiastic about engaging with CGD projects**, which presents civil society with an important opportunity. Therefore, while the road is long and contains many barriers, the following actions, if promoted and applied by civil society actors (and cooperative governments), can help lead to the constructive dialogue and a collaborative approach required to improve the effectiveness and sustainability of CGD projects:

- Establish good relations with local and national governments to promote mutual trust
- Establish solid partnerships with other civil society actors
- Seek international support
- Open you work!

To read the full report, click **here**.

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.

