



**Commission on the Status of Women  
Sixty-third Session**

**Women's empowerment and the link to sustainable  
development**

**INTERACTIVE EXPERT PANEL**

**Women's empowerment and the link to sustainable  
development – the data challenge and opportunity**

**Potentials of using Citizen-Generated Data for addressing  
gender data gaps: Africa regional perspective**

by

Davis Adieno\*

Global Partnership for Sustainable Development Data

Monday, 18 March 2019

10:00 am – 1:00 pm

---

\* The views expressed in this paper are those of the author and do not necessarily represent those of the United Nations

## Background

According to DataShift<sup>1</sup> – an initiative CIVICUS World Alliance<sup>2</sup>, Citizen-Generated Data (CGD) describes data that people or organisations produce to directly monitor, demand or drive change on issues that affect them. CGD is data generated by citizens that falls outside the remit of official data or statistics. In most cases its production is initiated by citizens or non-state actors for example Civil Society Organisations (CSOs), Community Based Organisations (CBOs), and religious institutions through research, social audits, crowd-sourcing online platforms, mobile phone and SMS surveys, phone calls, reports, storytelling, social media, and community radio. CGD can be quantitative or qualitative, structured or unstructured, and open or closed. It comes in a number of formats, ranging from numerical data in spreadsheets to text, audio or photos. Typically, citizen-generated data is collected through a specific initiative that aims to have a positive social impact. For example, an initiative may directly want to address corruption, sexual harassment, service delivery, or environmental degradation. Anyone can set up a citizen-generated data initiative, including businesses (e.g. Premise<sup>3</sup>), governments (e.g. the Abra Community Employment and Development Program<sup>4</sup>) and international institutions like the UN (e.g. GeoTag X<sup>5</sup>).

CGD initiatives harness the power of collective intelligence to contribute in areas where key data is missing – like the health of our oceans. Plankton Portal<sup>6</sup>, which uses crowdsourcing to locate and classify different photos of plankton to help scientists better understand the function and health of the ocean from small to global scales. This can be used to directly measure progress on Sustainable Development Goals (SDGs) number 14 on life below water. CGD in most instances is produced for very specific purposes which may or may not complement official data or statistics. It's therefore useful in plugging data gaps or checking the accuracy and quality of government data being produced. In China for example, Float Beijing<sup>7</sup> gathered citizens in the city to build air quality sensors attached to kites that could produce an accurate, timely dataset on air quality in the city because the Chinese government wasn't publishing sufficient air quality information. Float Beijing could be used to monitor targets and indicators under SDG 11 on sustainable cities and communities. In other cases, citizens collect qualitative data to raise awareness of a topic that isn't getting enough attention from institutions – like HarassMap<sup>8</sup>, which collects experiences of sexual harassment in Egypt to raise awareness on this important issue and could be used to directly monitor progress on targets under SDG 5.

In the spirit of collective responsibility embodied in the 2030 Agenda we need to recognize that sometimes citizens are simply better placed to gather data on a particular topic. SDG 5 on "Achieving Gender Equality and the Empowerment of Women & Girls" sets target 5.4 to recognize and value unpaid care and domestic work through the provision of public services, infrastructure and social protection policies. This is an example of a target whose data can be more accurately collected through self-enumeration by citizens since it may require real-time

---

<sup>1</sup> <http://civicus.org/thedatashift>, accessed January 2017

<sup>2</sup> <http://www.civicus.org/>, accessed January 2017

<sup>3</sup> <http://civicus.org/thedatashift/wp-content/uploads/2016/05/case-study-premise.pdf>

<sup>4</sup> <http://civicus.org/thedatashift/wp-content/uploads/2015/07/case-study-ccagg.pdf>

<sup>5</sup> <http://civicus.org/thedatashift/wp-content/uploads/2016/05/case-study-geotag-x.pdf>

<sup>6</sup> <https://www.planktonportal.org/>

<sup>7</sup> <http://civicus.org/thedatashift/wp-content/uploads/2015/07/Float-Beijing-case-study.pdf>

<sup>8</sup> <http://harassmap.org/en/>

data collection within the home. Check My School<sup>9</sup> is an example of a participatory public education monitoring program in the Philippines that enables parents to send in feedback about schools via SMS, twitter and other media – also connecting them with the Department of Education to help inform policy making. Furthermore, according to DataShift, a broad range of approaches and technologies can be used for collecting citizen-generated data. These include:

- Undertaking new surveys (e.g. to measure literacy<sup>10</sup>) whether via more traditional household questionnaire formats or new tech-dependent formats like SMS and radio feedback mechanisms (e.g. to measure attitudes to disease prevention<sup>11</sup>);
- Deploying monitoring equipment – including custom devices (e.g. to map pollution<sup>12</sup>);
- Mapping with “drones” or GPS devices (e.g. to scrutinise land boundaries<sup>13</sup>);
- Combining multiple existing databases (e.g. to count migrant deaths<sup>14</sup>);
- Scraping and aggregating data from official sources (e.g. to monitor official pardons<sup>15</sup>);
- Cross-referencing official, news and social media sources (e.g. to count police killings<sup>16</sup>; to assess public service delivery<sup>17</sup>);
- Creating crowdsourcing mechanism in order to collect individual stories and reports from citizens and civil society groups (e.g. to better measure marine debris<sup>18</sup>; to monitor economic conditions<sup>19</sup>);
- Developing micro-tasking platforms to utilise online assistance in performing tasks that require human cognition (e.g. to monitor deforestation<sup>20</sup>; to assist in disaster relief efforts<sup>21</sup>);
- Social or community-based auditing which empower citizens to undertake their own inquiries into issues that affect their lives (e.g. National Taxpayers Association social auditing<sup>22</sup> of Constituency Development Funds in Kenya or Community Based Auditing to improve natural resource planning and management<sup>23</sup>).
- A combination of more than one of these approaches (e.g. to map the supply of and access to water at district level<sup>24</sup>);
- DataShift’s project with the Open Institute<sup>25</sup> in localising the SDGs at the community level in Lanet Umoja Location, Nakuru County, in Kenya with Chief Francis Kariuki<sup>26</sup>, demonstrated (at a small scale) the sort of effort needed to reach everyone, including the most marginalised, to empower them through data literacy and CGD collection and effective use.

## Renewed global interest in complementary sources of data to track progress on SDGs

---

<sup>9</sup> <http://www.checkmyschool.org/>

<sup>10</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/uwezo/>

<sup>11</sup> <http://www.africasvoices.org/case-studies/unicef-somalia/>

<sup>12</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/citizen-sense/>

<sup>13</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/community-drones/>

<sup>14</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/the-migrants-files/>

<sup>15</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/who-is-pardoned/>

<sup>16</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/the-counted/>

<sup>17</sup> <http://civicus.org/thedatashift/wp-content/uploads/2016/05/case-study-pulse-lab.pdf>

<sup>18</sup> <http://civicus.org/thedatashift/wp-content/uploads/2015/07/case-study-dive-against-debris.pdf>

<sup>19</sup> <http://civicus.org/thedatashift/wp-content/uploads/2016/05/case-study-premise.pdf>

<sup>20</sup> <http://civicus.org/thedatashift/wp-content/uploads/2016/01/case-study-forestwatchers.pdf>

<sup>21</sup> <http://civicus.org/thedatashift/wp-content/uploads/2016/05/case-study-geotag-x.pdf>

<sup>22</sup> <http://www.fiscaltransparency.net/resourcesfiles/files/20150831126.pdf>

<sup>23</sup> <http://www.sciencedirect.com/science/article/pii/S0016328709002055>

<sup>24</sup> <http://civicus.org/thedatashift/learning-zone/research/changing-what-counts/who-actually-has-access-to-water/>

<sup>25</sup> <https://openinstitute.com/what-we-know-about-cdg/>

<sup>26</sup> <https://twitter.com/Chiefkariuki>

Following the launch of the Sustainable Development Goals (SDGs) in September 2015, the pressure to establish whether or not progress is being made has piled in recent times. Timely and relevant data that can help decision makers and other leaders better understand the current state of play for the 2030 Agenda has now come in to sharp focus. Real-time monitoring of the SDGs is top of the agenda of key global, regional and national leaders, and other actors. This has renewed interest in other sources of data like satellite imagery, mobile data, and Citizen-Generated Data (CGD) that can adduce more timely insights, complement official statistics, and plug data gaps in ways that help change-makers across the world achieve faster progress.

### **Gender data gaps and the limitations of Citizen-Generated Data**

According to a recent report prepared by Open Data Watch with support from Data2x "Bridging the Gap: Mapping Gender Data Availability in Africa", out of the 104 gender relevant indicators reviewed in 15 Sub-Saharan African countries, 48 percent of gender-relevant indicators are missing or lack sex disaggregated data. In international databases, 22 percent of the indicators lack any sex-disaggregation and 26 percent are missing data entirely. In national databases there are more missing observations (35 percent) but a smaller proportion – 13 percent – that lack sex-disaggregation. The report further notes that the persistence of large gaps in both international and national databases points to the need for a coordinated effort to improve data collection and adopt common standards for the compilation of indicators.

At the country level in Africa, many organizations working on gender related issues are doing an amazing, often invisible work at the community. Most of these organizations lack systematic ways of collecting and managing data. Often, they are overwhelmed by the sheer magnitude of the issues they are trying to address, exuberated by deep-seated cultural and societal beliefs that mount significant barriers in their work. The importance of data is also often obscured by the need to find practical solutions to real life challenges that sporadically emerge at that level. The overall impact of this however, is that we don't get a comprehensive or realistic picture of the magnitude of the issues these organizations are trying to address, for example GBV or FGM in the communities they serve. There are many undocumented or poorly documented incidences because community-based organizations (working where it hurts the most) lack a systematic time series record of the longer-term impact of these practices. In the end, they don't catch the attention of policy makers or decision-makers. And the fact that they can't demonstrate this impact in concrete terms means they find it difficult to rally political or public support behind the issues.

There're certain cultural sensitivities that comes with the collection of sex disaggregated CGD. Cultural biases in some instances make it difficult to ask respondents certain types of gender-related questions. For example, in communities where wife-battering is common practice but illegal, any research questions around the issue will often elicit hostility. Furthermore, if cultural perceptions by both men and women are such that GBV is some form of generally "acceptable discipline" it makes it all harder to tackle the issue. These cultural nuances are often missed in official statistics, but can be very well surfaced by CGD.

Context specific nature of certain SDGs targets means that organizations working on SDG 5 for example have to target their interventions and collect relevant data that speaks to the needs of those communities. I recall in earlier research in Kenya it was brought to my attention that FGM was not practiced at all in one of the communities, but GBV was prevalent. Respondents therefore felt that it would be better for me to focus on GBV for an in-depth study, rather than

cover everything and gloss over what they considered more pertinent for their circumstances (GBV).

Sharing of data remains a big challenge – even among non-state actors. In my experience most CSOs opt to share their data with donors, and only if that is a requirement. Many of them also contract out research or data collection to consultants whom they don't impress upon to deliver both raw data and the final report. They therefore have no data to share in the first instance. This is primarily driven by lack of internal capacity to generate or manage data or low regard for raw data. The situation has improved in recent times with more organizations appreciated the centrality of data in the delivery of the SDGs.

One of the counter-arguments by statisticians about CGD is that does not in most instances adhere to the Fundamental Principles of Official Statistics, therefore posing challenges to policy makers that may want to use it. DataShift noted that despite the immense potential of citizen-generated data, a number of both real and perceived challenges exist regarding its collection and use. These include limited coverage and potential to achieve scale; lack of “representativity” because CGD initiatives will often not be a representative sample of the whole population; limited credibility in the absence of standards or agreed good practices for CGD collection and use that compromise quality and reliability of the data; challenges with comparability given limited or no easy methods for comparing CGD collected by different actors; and usability with many projects failing to adequately consider how the information and datasets they produce will actually be used.

### **Opportunities for CGD to fill gender data gaps and improve understanding on key aspects of girls' and women's lives?**

A report dubbed “Advancing sustainability together? Citizen-generated data and the Sustainable Development Goals,<sup>27</sup>” commissioned by the Global Partnership for Sustainable Development Data (GPSDD) and developed by a team of experts concluded that dealing with data is usually much more than ‘just producing’ data. CGD initiatives open up new types of relationships between individuals, civil society and public institutions. This includes local development and educational programmes, community outreach, and collaborative strategies for monitoring, auditing, planning and decision-making. These relationship aspects of CGD are important when dealing with SDG 5 because they connect humans to build relationships that go beyond data production and use processes. The report further outlines other CGD benefits as:

1. CGD initiatives can help in gathering data in regions otherwise not reachable. Some CGD approaches may provide updated and detailed data at lower costs and faster than official data collections. CGD can highlight populations and the situations of those being left behind in marginalized and hard to reach areas in order to inform policy decisions and targeted resource allocation.
2. Beyond filling data gaps, official measurements can be expanded, complemented, or cross-verified. This includes pattern and trend identification and the creation of baseline indicators for further research. CGD can help governments detect anomalies, test the accuracy of existing monitoring processes, understand contextual factors, and initiate their own follow-up data collections.

---

<sup>27</sup> <http://www.data4sdgs.org/resources/advancing-sustainability-together-citizen-generated-data-and-sustainable-development>

3. CGD can inform several actions to achieve the SDGs. Beyond education, community engagement and community-based problem solving, this includes baseline research, planning and strategy development, allocation and coordination of public and private programs, as well as improvement to public services.
4. CGD must be 'good enough' for different (and varying) purposes. Governments already develop pragmatic ways to negotiate and assess the usefulness of data for a specific task. CGD may be particularly useful when agencies have a clear remit or responsibility to manage a problem.
5. Data quality can be comparable to official data collections, provided tasks are sufficiently easy to conduct, tool quality is high enough, and sufficient training, resources and quality assurance are provided.

The other distinct advantages of CGD especially for gender related issues include:

1. It offers the opportunity to ordinary citizens, untrained in formal data and statistics methodologies to be creative and flexible in development and governance interventions; often better understanding and responding to people's immediate and rapidly evolving needs. If a community is facing the menace of FGM citizens have the power and opportunity to address the issue among themselves.
2. The contextualised nature of CGD means it can yield highly disaggregated data to the lowest levels of the communities on SDG 5 targets and indicators, highlighting specific needs and priorities for government and other stakeholder's interventions.
3. CGD empowers citizens to better understand their development and governance context, therefore have the ability to hold their leaders to account for the delivery of promises made under national development plans and the SDGs. By surfacing nuances around SDG 5 it can empower organizations to formulate better strategies for tackling issues within communities.
4. CGD is often gathered on themes and topics that matter to citizens, potentially flagging up issues of social injustice, gender biases, economic inequality or environmental degradation that might otherwise be missed by official statistics.
5. CGD is actively given by citizens, providing direct representations of their perspectives and an alternative to datasets collected by governments or international institutions.
6. CGD creates greater ownership of the development discourse among citizens. With a greater sense of ownership citizens can identify their priority targets and indicators meaningfully contributing to the delivery of the SDGs. It is easier to engage communities to change their circumstances using the data they have collected themselves.
7. CGD can help achieve scale on ideas and interventions that work, especially through the testing of new, innovative ideas, and developing practical models at various levels of development. If a problem evolves and assumes a new dimension, for example harassment, organizations working on this would have to quickly adopt their approaches and methodology to obtain the relevant evidence and act quickly to address it.

### **In conclusion**

All in all, I am passionate about the renewed focus on data under the 2030 Agenda. I however feel we urgently need a more energized focus on data products and their insights to inform evidence-based decisions, but more importantly action. On gender related issues directly affecting communities, the statistical significance of Citizen-Generated Data is just as important, or even more important than that of official statistics in tackling these issues in a timely manner.

If it is a matter of life and death, ultimately the life of one woman or one girl in a remote village in a Sub-Saharan African country is more important than the statistical significance of that number when they are gone because we stood by and took no action to prevent their death in the first instance. Let's humanize the data and inspire action - quick!