TOWARDS ENGAGING CIVIL SOCIETY ORGANISATIONS IN THE OPEN DATA AGENDA
A CASE STUDY IN THE HEALTH SECTOR IN TANZANIA

RESEARCH PAPER
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2 Problem statement</td>
<td>4</td>
</tr>
<tr>
<td>3 Objective</td>
<td>4</td>
</tr>
<tr>
<td>Main objectives</td>
<td>4</td>
</tr>
<tr>
<td>Specific objectives</td>
<td>4</td>
</tr>
<tr>
<td>4 Methodology</td>
<td>4</td>
</tr>
<tr>
<td>Sampling</td>
<td>4</td>
</tr>
<tr>
<td>Data collection</td>
<td>4</td>
</tr>
<tr>
<td>5 Findings and discussion</td>
<td>5</td>
</tr>
<tr>
<td>Data storage</td>
<td>5</td>
</tr>
<tr>
<td>Data collection</td>
<td>5</td>
</tr>
<tr>
<td>Data quality</td>
<td>6</td>
</tr>
<tr>
<td>Data sharing and reuse</td>
<td>6</td>
</tr>
<tr>
<td>Data policies</td>
<td>7</td>
</tr>
<tr>
<td>6 Conclusions and recommendations</td>
<td>7</td>
</tr>
<tr>
<td>References</td>
<td>9</td>
</tr>
<tr>
<td>Appendix A: Questionnaire</td>
<td>10</td>
</tr>
</tbody>
</table>

---

**Authors**
Mahadia Tunga and Joseph Mushi

**Publication date**
October 2016

**Published by**
World Wide Web Foundation

**Licence**
Published under a Creative Commons Attribution International 4.0 Licence

**Suggested citation**

**Cover photo credit**

---

This work was supported by the World Wide Web Foundation as part of the Open Data for Development (OD4D) network and carried out with financial support from the UK Government’s Department for International Development (DFID) and the International Development Research Centre (IDRC), Canada. The views expressed in this work are those of the creators and do not necessarily represent those of the Web Foundation, the DFID and the IDRC or its Board of Governors.
Abstract

Despite open data initiatives in the public sector, little has been done in the private sector and civil society sectors. This study analysed open data initiatives in civil society organisations (CSOs) in Tanzania with a special focus on the challenges that hinder the participation of health-based CSOs in open data initiatives. The health sector was selected because it is one of the priority sectors identified in the Tanzanian’s Open Government Partnership Implementation Plan. CSOs in the health sector have been collecting and generating vast amounts of data. The most common method of data collection is in the form of routine data for the health management information system (HMIS) collected by the focal person at the level of health facility or district. In most cases, there are therefore two different sources of data often resulting in conflicting information.

Semi-structured interviews were conducted with CSOs in the health sector. Interviews were complemented with focus group discussions with key stakeholders and with field visits. It was found that vast amounts of data have largely not been published and the few datasets that have been published are in PDF and other non-machine readable formats, which limit the usability and manipulation of such datasets in other analytical studies. The sharing of data amongst CSOs is not practised; consequently there are unnecessary duplications and inconsistencies which compromise the quality of the data and increase the cost and time of conducting field studies. Health CSOs’ tendency to withhold data has been influenced by the competition to secure donor funding, poor quality of data, inadequate support from donors, and the lack of necessary knowledge and facilities among CSOs to support open data initiatives. The paper concludes with a set of recommendations to improve the publication of open data among health CSOs in Tanzania.

1 Introduction

Governments across the world are actively engaging in the open data struggle (OGP 2014). The major concern has been information provision by public sectors and citizen participation with the common objective being to improve the monitoring of government activities and service delivery. Several CSOs across the country – such as Hakielimu and Twaweza – have been steering the struggle towards open data in public institutions to foster government accountability (Bussiek 2015). For years, the issue of open data was seen as a matter for governments and their public institutions alone (Kikweke 2015), with the private sector and CSOs being left out. As a consequence, little has been done to exert pressure on the private sector and CSOs to engage with open data. At the same time, the government of Tanzania has recently enacted a new statistics act restricting the publication of data without government sanction (NBS 2015). This new act challenges the participation of CSOs in the open data agenda, and is said to deny the right to access information, which is contrary to ‘African statistics shall not be made inaccessible in any way whatsoever’ (NBS 2015).

In Tanzania there are thousands of CSOs working in the education and health sectors. In the health sector, CSOs play critical and diverse roles related to healthcare delivery and health needs. However, these CSOs have been working in silos and thus fail to adequately bridge the gap in facilitating access to equitable health services (Macharia, Nouwen, Koome, Makhulo, Ndungu & Fall 2013).

CSOs working in the health sector collect routine data through the HMIS focal person at the district or facility (local) level. HMIS is a process whereby health data are recorded, stored and processed for policy-making, planning, implementation and evaluation of health programmes (Nyamtema 2010). There are two forms of HMIS – manual or electronic. With manual HMIS, data collection is done through recording daily patients’ records in counter books by health facility employees. At the end of the month, the records entered in counter books are summarised in HMIS forms and submitted to the HMIS focal person at the district, who then enters the collected data in the electronic HMIS system called the district health information system version 2 (DHIS2).

In Tanzania’s health system, there are several kinds of health facilities – dispensaries, health centres and district, regional and referral hospitals – which are located all over the country (though the coverage is not enough to meet demand). Health facilities are either managed by the government and/or by private institutions.
2 Problem statement

In Tanzania and the rest of sub-Saharan countries, there are many CSOs that are actively engaged in health sector activities who have been collecting and generating data for quite a long time (Macharia et al. 2013). They are therefore an important source of data. However, the data collected and produced by these CSOs have largely not been published. The few datasets that have been published are in PDF and other non-machine readable formats, limiting the usability and manipulation of such datasets in other analytical studies.

Moreover, CSOs do not share data among themselves and, to large extent, do not use datasets published by other CSOs or by the government. Many CSOs have conducted field studies to collect data that have already been collected by other CSOs. This has resulted in unnecessary duplication and inconsistency, compromising the quality of the data and increasing the cost and time of conducting field studies.

3 Objective

MAIN OBJECTIVES

The main objective of the research was to determine factors inhibiting CSOs in the health sector to share and reuse datasets published by other CSOs in order to improve the open data ecosystem in Tanzania.

SPECIFIC OBJECTIVES

The specific objectives of the research were to:

- Analyse the current status of data collected by CSOs in the health sector;
- Study the procedures, standards and policies governing data collection and dissemination;
- Determine challenges associated with sharing and reusing datasets among CSOs in the health sector;
- Suggest ways to improve sharing and reusing datasets among various CSOs in the health sector.

4 Methodology

The study was conducted from November 2015 to January 2016. Combinations of methods were used to collect data. This approach was taken to minimise the risk of data-error during the collection process. The methods used included a review of secondary data; semi-structured interviews; organised focus group discussions with key informants; and field visits.

The term CSO has acquired various definitions among different actors. However, for the purpose of this study, CSO includes non-profit and non-governmental organisations outside of the public sector in which people organise themselves to pursue shared interests in the public domain (Tomlinson 2013).

SAMPLING

A preliminary review of the literature was carried out to inform the selection of CSOs for the study. The main purpose of the preliminary review of documents was to orientate the research team about the CSOs operating in the health sector in Tanzania. For reasons of budget, the study limited its selection to CSOs operating in Dar es Salaam, the capital city of the Republic of Tanzania. The criteria for the selection of these organisations were that the selected CSOs:

- Are involved in the health sector in Tanzania;
- Have an operational office in Dar es Salaam;
- Have at least five years’ experience of working as a CSO in the health sector.

On the basis of the above criteria, the study selected to interview members of staff from ten CSOs, namely: Ifakara Health Institute (IHI), Ariel Glaser Pediatric AIDS Healthcare Initiative (AGPAHI), Benjamin William Mkapa HIV/AIDS Foundation (BMAF), Christian Social Services Centre (CSSC), National Muslim Council of Tanzania (BAKWATA), Management and Development for Health (MDH), John Snow INC Tanzania (JSI), Elizabeth Glaser Pediatric Aids Foundation (EGPAF), Aids Relief and Tanzania Youth Alliance (TAYOA).

DATA COLLECTION

After the document review and the selection of CSOs, the study conducted semi-structured interviews based on a prepared questionnaire (see Appendix A). Two staff
members from each of the ten CSOs selected, were interviewed. To verify information gathered during semi-structured interviews, the study organised focus group discussions with key informants as well as field visits. Two sites were visited – Mwanyamala Hospital and Kinondoni municipal council – in order to observe the routine data collected by the HMIS focal person.

5 Findings and discussion

This section presents the research results and the analysis of those results.

DATA STORAGE

Respondents indicated that 60% of CSOs in the health sector store data on local personal computers due to the lack of a common repository for field data within the respective CSO (Figure 5.1). In these CSOs, data are stored on the focal person’s computer where data is stored until the release, acceptance and dissemination of an intended report for which the data were collected. Once the report has been released, data remain under the custody of the focal person, who could be the data analyst, from which its survival depends on the wishes of the respective custodian. This observation was supported by data depicted in Figure 5.2, which show that nearly 80% of CSOs only keep data for the period of a project’s lifespan. These findings indicate that most CSOs are performance based; they exhaust a lot of effort on preparing reports to submit to a donor rather than on data storage for future reuse or sharing.

DATA COLLECTION

Based on the semi-structured interviews, the methods used by CSOs for data collection were positively correlated \((r=1)\) with the type of data collected. The responses of interviewees about the type of data collected by their respective CSOs indicated that 68.4% of CSOs routinely collect health-related data whereas 31.6% collect survey data (Figure 5.3).

This observation was supported by the focus group discussions. It was observed that the majority of CSOs in the health sector, through the outreach department and/or monitoring and evaluation personnel, collect data routinely from a health facility or district HMIS focal persons. A facility focal person is usually a local worker in a health facility, who usually collects data daily by writing patient records in counter books. Counter books are a manual means of collecting health data in which, at end of the month, the health facility personnel summarise the data into HMIS forms and submit them to the HMIS focal person at the district level. At the district authority, the focal personnel assemble monthly reports from all facilities within respective districts and capture the data in the DHIS2 database. As a result, some CSOs collect data directly at the facility level while others collect at the district level. Accordingly, in most cases, these two different approaches result in conflicting information.
DATA QUALITY

During the focus group discussions, it emerged that the quality assurance exercise is another area in which there is duplication of effort. According to participants, CSOs in the health sector do not normally employ data collectors for routine data collection at health facilities – instead they use data collected by health facilities. In order to assure the quality of the data collected by health facilities, CSOs perform analyses and if a gap is identified in the collected data, they then conduct a supportive supervision exercise. According to participants, health authorities at the district level are responsible for assigning areas for conducting supportive supervision to the CSOs. But sometimes, due to poor coordination, more than one CSO may be assigned in one district to perform supportive supervision for similar activities. For instance, in 2015, the CSOs Tabasamu and EGPAF deployed supportive supervision in the Nzega District. Both CSOs were assigned to supervise the same district officer recording reproductive and child health data. If the list of CSOs and their respective areas of supervision were published, then such duplication of effort and resources would not have happened. Instead the district authority would have been notified and one of the CSOs could have been allocated to other unsupervised areas.

DATA SHARING AND REUSE

Types of published data

CSO interviewees were asked about issues related to data-sharing practices within and between CSOs, as well as between CSOs and the public at large. Based on the information provided by the CSO respondents in the interviews and focus group discussions, the majority of data shared with the public by CSOs were those related to success stories in relation to their activities, which mostly centre on awards from competitions and donor funding. Of the CSO respondents, 73.5% indicated that they publish repackaged data as evidence of their successes while 26.5% publish aggregated data as evidence of service delivery (Figure 5.4). None of the CSOs interviewed publish financial or raw data.

Characteristics of published data

It was observed during the interviews that financial and raw data are not published for various reasons. These include but are not limited to donors’ policies placing restrictions on the sharing of data related to project funding and/or releasing raw data to the public. In this regard, only CSOs that are funded by the same donor are able to share data. In addition, other factors contributing to the limited publication of data were: competition among CSOs to secure funding; failure to follow appropriate procedures set by the National Board of Statistics and the health ministry; the poor quality of the data; and the sensitivity of health data containing personal data that could not be shared.
Towards engaging civil society organisations in the open data agenda

Water points were both operational while nearby health facilities lacked the same service. These instances of the duplication of service provision could possibly have been avoided if the CSOs had been accessible to each other and actively shared their data.

Based on the interviews and the focus group discussions with the CSOs, the main reasons as to why the majority of CSOs do not favour the publication of data for reuse by other CSOs concern donors’ policies and funding competition. Other reasons include limited access due to lack of a central point for data exploration; unreliable sources because most data provided by external sources do not provide full details on the methodology used in data collection; and the difficulty in assessing the quality of the data published by other CSOs.

DATA POLICIES

Most CSOs operating in the country lack policies for data collection and dissemination. Close to 90% of respondents indicated that their CSO does not have or use established standards and policies to govern data collection and dissemination. Rather, they adhere to the standards and regulations of donor agencies. As a result, policies within CSOs change depending on the nature of project and the requirements identified when applying for project funding. The application of non-harmonious donor policies has largely affected the publication and sharing of data, such that only organisations supported by the same donor are likely to share data for further analysis and publication.

6 Conclusions and recommendations

The study has revealed a persistent lack of necessary knowledge and facilities among CSOs for supporting open data initiatives. Although CSOs have invested in the collection of data in the health sector, the dissemination of raw data as well as making analysed records accessible to the public has been negligible. When data is made available, the data are typically made available in non-machine readable formats such as portable document format (PDF). Based on findings of this study, the following recommendations are made for the improvement of data sharing within and between CSOs, and between CSOs and the public at large.

National standards. In order to ensure the smooth exchange of data, there is a need to establish standardised tools for health data reporting in the country in such a way that data from different sources can be merged and used for further analysis.

Central database. The findings of this study revealed that the majority of CSOs in Tanzania lack the necessary capacity for data storage. This results in CSOs storing data temporarily and not for a period longer than a project’s lifespan. Therefore, the provision of a central database as a repository for CSOs’ health data would contribute to motivating CSOs to store their data beyond a project’s lifespan.

Moreover, a central database will provide room for data sharing among CSOs and minimise the duplication of data collection at different levels. The Millennium Challenge Corporation (MCC), in collaboration with the US President’s Emergency Plan for AIDS Relief (PEPFAR), have recently launched a project to set up a data collaboration lab which will provide centralised access to data published by various sectors in the country. The project will start with the health sector in Tanzania. The project aims not only to provide space for data storage but also to build capacity in data manipulation, analysis and sharing.

Improved data quality. The majority of CSOs argued that data collected by other CSOs are unreliable because most of them do not include details on the methodology used in the collection of the data, which, in turn, introduces issues related to the quality of the data published by CSOs. Therefore detailed methodological and other information should accompany published data so that other CSOs can reuse the data. In addition, there is a need to improve the level of detail present in datasets for other CSOs to utilise data for further analysis. With the current practice of publishing aggregated data, it is difficult to employ such data for other uses.

Strict allocation of areas of work among CSOs. Currently there is a distribution of areas of work among CSOs in the health sector in Tanzania. However, there is no strict follow-up on the same. As a result, the duplication of effort still occurs in some areas of work. Therefore, there is a need to tighten measures so that CSOs are restricted to allocated areas for the purpose of avoiding duplication as well as increasing the coverage of health data collection for effective sharing and reusing.
Discourage donor restrictions. Efforts should be made to discourage restrictions imposed by donors on the sharing of data between CSOs, and between CSOs and the public. Every project should encourage the creation of an environment for data sharing among CSOs. This can be done by providing the tools and skills currently lacking in CSOs in Tanzania, as well as by formulating policies that will promote data sharing among different CSOs.

CSOs should not only rely on donors to amend their data sharing policies, but should also take on the responsibility to coordinate among themselves towards producing and adhering to common standards and data sharing policies within the health sector. By doing so, CSOs will not only create a more harmonious data sharing environment, but they will bolster their collective bargaining power in negotiating for more open data sharing policies from donor agencies and government.

National policies and regulations governing data collection and dissemination. In 2015, the government of Tanzania enacted a statistics law that has implications for setting limits on the dissemination of data. The act argues that the non-governmental sector should either involve a government official from the National Bureau of Statistics during data collection or submit the reports to the Bureau before publication (URT 2015). This law limits organisations from publishing and exchanging information. The statistic act needs to be challenged by different activists in order to provide the necessary freedom for the collection and dissemination of data.

On the other hand, the government of Tanzania is in the process of establishing an open data policy. The policy objectives include creating a platform for sharing technical expertise and experiences among actors in the private sector, thus encouraging the private sector to open up their data to enhance collaboration in research and development (URT 2016), including between CSOs.
REFERENCES


Appendix A: Questionnaire

TOWARDS ENGAGING CIVIL SOCIETY ORGANIZATIONS (CSOS) IN OPEN DATA AGENDA

A Case of Selected CSOs in Health Sector in Tanzania

• This questionnaire aims to collect data that will contribute in finding a solution on how to engage CSOs in open data agenda. Your willingness to participate is very much appreciated.

• Your responses will be treated in confidence and used for the purposes of this study only.

• Completion of this questionnaire is completely optional.

For further information contact mahadiatunga@gmail.com +255713 190018.

Full name (optional) ...........................................................................................................................................................

Organization (optional) ...........................................................................................................................................................

1.0 Current status of data collected by CSOs

i. What kind of data do you collect?
........................................................................................................................................................................................

ii. How do you store your data?
........................................................................................................................................................................................

iii. For how long do you keep the stored data?
........................................................................................................................................................................................

iv. What are your sources of data?
........................................................................................................................................................................................

v. How do you monitor quality of data?
........................................................................................................................................................................................
2.0 Challenges associated with sharing and reusing datasets among the CSOs in health sector

i. Do you publish data to the public?

ii. What kind of data do you publish?

iii. What kind of information do you withhold? Why?

iv. Do you share unpublished data with other CSOs?

v. In what form do you share data?

vi. Do you use data from other CSOs?

vii. Do you use data from national open data portal or government agencies like NBS?

viii. Why are you not using data from other agencies?

3.0 Procedures, standards and policies governing data collection and dissemination

i. What methodology do you use for data collection?

ii. Do you have policies governing data collection? Provide details

iii. Do you have policies governing data dissemination? Provide details

iv. Do you use any open data license when publishing data?
4.0 Ways that can be used to improve sharing and reuse of datasets among various CSOs in health sector

i. ...........................................................................................................................................................................

ii. ...........................................................................................................................................................................

iii. ...........................................................................................................................................................................

iv. ...........................................................................................................................................................................

v. ...........................................................................................................................................................................

Additional notes

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................

...........................................................................................................................................................................................
For more information on this project, visit our website at labs.webfoundation.org or get in touch with us directly at contact@webfoundation.org. Other related media are also available in the resources section of our website.

If you want to learn more about the project, email us at contact@webfoundation.org.

If you want to look at other open data projects, see www.webfoundation.org.

If you want to give it a try and implement the approach – with or without us – let's talk!