

MULTIDIMENSIONAL OPEN DATA COLLECTION TOOL



Time Needed:

A few hours



Materials Needed:

Set of data needs that have been identified by community workshops



Difficulty Level:

Moderate



Tool Creator:

Khairil Yusof

Version:

1.0, 30 April 2017

TOOL PURPOSE:

This tool will help you define, match and expand community identified open data against non-government organisations (NGO), government and industry taxonomies. This will help actors to align the open data needs of their communities with the goals of other stakeholders to increase collaboration, support/funding and impact for their open data project.

For example, this [social audit data from urban poor household survey¹](#) covers number of children and age of children, and overlaps with some [UNESCO data²](#), but is missing key data points that might have improved the dataset so that it would be more complete and could be used by UNESCO.

Where data is not provided by government, it is important to use alternative data that is up to international standards.

Similarly in the same survey, there were questions on cost of Internet and mobile phones. If the question was phrased to get data slightly following Web Foundation's [Digital Gender Gap Toolkit³](#) and Alliance for Affordable Internet (A4AI – a4ai.org), it would have improved data collection on gender access and affordability of Internet that could be compared against other countries or against other communities.

PREREQUISITES:

To successfully make use of this, you need to:

- Be someone familiar with programme work and tools such as logical frameworks. Logical Framework Approach is a method developed for USAID. It is another tool that helps an NGO programme officer map his/her activities and outputs to his/her overall goals and short term outcomes. He/she will be able to use a similar approach to initially map out goals of international development agencies to his/her own organisation, and then additionally map out overlapping data requirements.
- Somebody with community organizing skills - or the capacity to build and manage relationships among different stakeholders.

MULTIDIMENSIONAL OPEN DATA COLLECTION TOOL

- Once you have your framework and what data you need to get, you need to have some skills in data gathering (e.g. interviews, focus groups) to be able to locate the data assets needed and the ones holding them.
- Additionally the focal point must ideally be familiar with basic structured data. For example in this case, understanding that for household survey, individual data can be asked separately as separate survey form and linked to unique identifier such as house number.

Additionally, this tool works best when you already know the community data needs, but are not familiar with sources of open data and or potential partners or stakeholders, and how to map community data needs with that of partners and stakeholder goals and data needs.

WHO TO INVOLVE:

- Community stakeholders
- Local & international NGOs identified when using the tool
- Local government agencies identified/matched when using tool

STEPS TO TAKE:

1. Hold community workshops to identify open data needs of communities. This can be undertaken through informal conversations, or through organised workshops. A sample of a tool in identifying data needs can be found on pages 1 and 2 of the UK Data Service Research Question to Exploratory Data Analysis⁴. A sample of a data needs identification workshop design can be found in our Data Prioritisation Workshop methods⁵.
2. Match as many of the data needs against sector taxonomies of the tool, and expand set of open data points that could or should also be collected as part of this project.

Simply put, a taxonomy is a manner of classifying and arranging data that are otherwise meaningless and random when treated individually. Examples

MULTIDIMENSIONAL OPEN DATA COLLECTION TOOL

are found below:

Use case: Schools

Data needs: textbooks, classrooms(facilities/infrastructure), chairs/desks, teachers

Use case: Urban poor

Data needs: Household income, number and age of children, single parents, employment

Some sectors however have already a built data taxonomy (see for example the World Health Organization for health data, and UNICEF for education).

See the guide to developing taxonomies for effective data management⁶ for an example of a process in building a taxonomy. Another example of a taxonomy, this time in the agricultural sector, can be found in Tapipedia.org⁷.

3. Identify partners from which data that could be collected, and work with the partners in understanding the communities to see what open data they have and or support they can provide.

A community organiser or focal point will be able to identify the best methods and sources of data that matches the needs identified above. It could be primary data collection (e.g. a door to door survey) or sourced out through secondary sources (e.g. may be from data sourced from other NGOs identified in the logical framework). For example, in your framework, you might have identified UNESCO as an institution having overlapping goals with yours. They may already have the data regarding the data needs that are related to your goals as well.

A sample data mapping tool can be found on page 5 of UNGP's Data Asset Facilitation Tools⁸.

4. Finalise and improve the data sets that are feasible to be collected in addition to the primary data sets already existing at the level of communities.

MULTIDIMENSIONAL OPEN DATA COLLECTION TOOL

5. Prepare project proposal / document that details:
 - The data needs/data points
 - Stakeholders who collect/produce/aggregate this data
 - Primary and secondary sources
 - Methods of data collection.
6. Implement the plan above and explore data availability as well as openness. Help improve taxonomy of toolset for new matches, missing sectors and partners.

RESOURCES:

1. Source: http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS
2. Source: http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS
3. Source: http://webfoundation.org/docs/2016/12/WRO-Digital-Gender-Gap-Audit_Toolkit.pdf
4. You can find UK Data Service's Research Question to Exploratory Data Analysis at <https://www.ukdataservice.ac.uk/media/455319/pracdatawishlist.pdf>
5. You can find our workshop design for Data Prioritisation at <http://bit.ly/2pLylt>.
6. A guide to developing taxonomies for effective data management could be found at <http://www.computerweekly.com/feature/A-guide-to-developing-taxonomies-for-effective-data-management>.
7. An example taxonomy in the agricultural sector could be seen at <http://www.tapipedia.org/>.
8. UNGP's Data Asset Facilitation Tools are accessible at <http://www.unglobalpulse.org/news/data-asset-facilitation-tools>.

MULTIDIMENSIONAL OPEN DATA COLLECTION TOOL

9. Multidimensional Poverty Index: <http://hdr.undp.org/en/content/multidimensional-poverty-index-mpi>.
10. UNICEF data: <https://data.unicef.org/>.
11. UNODC data: <https://www.unodc.org/unodc/en/data-and-analysis/>.
12. ILO resources: <http://www.ilo.org/inform/online-information-resources/>.
13. WHO data: <http://www.who.int/healthinfo/statistics/en/>.
14. OCDS Documentation: <http://standard.open-contracting.org/latest/en/>.
15. CoST data: <http://www.constructiontransparency.org/home>.

SOUTHEAST ASIA OPEN DATA INNOVATION WEEK

This tool was created by participants of the Southeast Asia Open Data Innovation Week event held last November 2016, as part of the Open Data Innovation Toolbox. There are more tools to check out and use on our website!

Visit <http://labs.webfoundation.org/projects-2/open-data-innovation-week/> to read about the project and see the other tools. For more information on the event, you can also get in touch with us by email at info@labs.webfoundation.org.