

Navigating the Politics of Open Data

A High-Level Political Forum Side Event

July 15, 2019

Hosted by the Individual Deprivation Measure Program¹, Open Data Watch, and UN Women



Background

In 1995, at the Fourth World Conference on Women, the **Beijing Declaration and Platform for Action** emphasised the importance of investing in sex disaggregated data, to provide data about the circumstances of women. It called on national and international statistical organisations to “collect gender and age-disaggregated data on poverty” and “examine the relationship of women’s unremunerated work to the incidence of and their vulnerability to poverty.” The Sustainable Development Agenda 2030, including the commitment to leave no one behind, has increased demand for disaggregated data to understand how people’s circumstances are affected by sex, age, disability, and other factors. Individual-level data is key to revealing which groups are experiencing what barriers, including multiple and overlapping barriers.



Photo: Alice Floyd/IWDA

This photo does not include IDM survey participants.

Disaggregated datasets, however, are only valuable to advocates and decision makers if they can access them. Discussions at the second UN World Data Forum in Dubai in October 2018 highlighted that making data available in a way that can be explored by users also supports transparency, and trust in data and in government. The openness of, and access to, these datasets is crucial to realizing the full social and economic benefits of the data and to leave no one behind. But access to datasets brings political and privacy challenges that must be managed. Producing and opening the types of quality datasets needed to accelerate the achievement of the Sustainable Development Goals (SDGs) will be facilitated by managing potential sensitivities arising from the data and highlighting the benefits of openness.

This High-level Political Forum side event, hosted by the Individual Deprivation Measure Program, Open Data Watch, and UN Women, is designed to enable sharing and documentation of emerging experience to inform and support work towards better, more accessible development data. Participants include individuals and representatives with experience of the challenges and opportunities of making data open and accessible, who can contribute to understanding how challenges are being navigated and opportunities maximized. The event will have some emphasis on data to fill gender data gaps. This document provides background on the important issues to be discussed at the event.

¹ A partnership between the Australian National University and the International Women’s Development Agency, funded by the Australian Government.

Create disaggregated datasets

The lack of gender-relevant data that can be disaggregated by sex, as well as other factors including age, disability and geographic location, limits our knowledge of the status and well-being of women and girls in countries around the world. A **Data2x study from 2014** noted that “Less than a third of countries disaggregate statistics by gender on informal employment, entrepreneurship, violence against women, and unpaid work.” In 2019, Data2X and Open Data Watch undertook an assessment of the availability of 104 gender-relevant indicators in 15 Sub-Saharan African countries. The resulting report **Bridging the Gap: Mapping Gender Data Availability in Africa** confirmed the paucity of sex disaggregated data, revealing that 48 percent of gender-relevant indicators are missing or lack sex-disaggregated data in the study countries at both international and national levels. Investments are needed to create these disaggregated data. New innovations in technology can help fill some data gaps. In other cases, new measurement approaches are needed to address data gaps or gender-blindness in available data. Making the case for better data, that measures what matters for gender equality and can be disaggregated, is key to improving the availability of quality data. One example of this is the **Individual Deprivation Measure (IDM)** a new, individual-level, gender-sensitive, measure of multidimensional poverty. The IDM collects primary data from all adult members of a household on 15 dimensions of life, including ones particularly relevant to the circumstances of women such as voice, time use and family planning. The resulting data provides insight into the uneven experience of deprivation both inside and outside the household – a feature unique to IDM data. The data can also be used to reveal the relationship between poverty and factors such as gender, age and disability, to inform priorities, support targeted action, and assess impact. Because IDM data makes it possible to see where deprivation and inequality is concentrated (which groups, what dimensions, where) and how factors overlap and intersect to deepen marginalisation, it is an example of both the relevance of disaggregated gender-relevant data for addressing inequality and poverty, and the potential sensitivities to be navigated.

Open up disaggregated datasets

Disaggregated datasets that are open and available to the public for use produce more societal value than closed datasets as they can be freely used to improve government programs and inform citizens. There are also economic benefits, and **research on the opening of Landsat satellite data** in the United States (US) estimates savings of between USD 350 and 436 million per year for the Federal Government, and the private sector. The 2018-19 **Open Data Inventory (ODIN)**, released in March 2019, provides the latest assessment of coverage and openness of official data and statistics in 178 countries. It shows an improving trend overall, with 11 countries increasing data openness by over 35 percent in two years. But as more disaggregated datasets become available, consistent political support and financing are needed to continue this progress. Changes to legal frameworks and access to information policies may also be needed to support open data policies. Incorporating open data in countries’ National Strategies for the Development of Statistics, and in the planning and implementation of SDG national reporting platforms, can signal the value of open data and provide a foundation for shifting the ‘default’ towards making data open. In a context where financing for statistics is constrained, making data open helps to maximise its value.

Provide access to data while protecting privacy

The granular level of information provided by disaggregated datasets to support efforts to leave no one behind also raises questions of privacy. While there are clear benefits of providing more granular insights, opening such data to the public widens access, and so requires particular assurances in relation to respondent privacy, anonymization, and data security, as is the case with microdata or administrative data. While most microdata providers protect survey respondents’ personal information by using a variety of **de-identification techniques**, even anonymized datasets can be de-anonymized if they have a large number of columns, attributes, and features can be joined with other datasets to re-identify participants,

as was done by two computer scientists during a **Data for Development Challenge**. There are further steps that data stewards can take to protect their datasets by sharing their data only with a trusted framework of users. Trusted user frameworks are used in scientific and other research-heavy fields for granting researchers access to otherwise closed data on the condition that they do not disclose any confidential material or data. There are also more technical solutions to this problem, such as the **Open Algorithms project (OPAL)** that ensures data privacy by only “sending the algorithms to the data,” so that people given access cannot see the data (which are kept safe by the agency housing them) but can still perform analyses on the data. The World Bank’s **Open Data Analysis Toolkit** has demonstrated that an approach that enables querying of datasets while the data itself remains inaccessible and secure can be achieved with large official datasets.

Navigate the politics and maintain trust in data

As the title of the session suggests, there are politics involved in opening and disaggregating data. Data are political in the sense that power is embedded within numbers. For example, the results of a census affect budget and resource allocations and the drawing of political districts. In other cases, disaggregated data can yield insights on minority populations that may support or negate official positions by Governments. In the case of gender data, feminism’s rallying call that ‘the personal is political’ encapsulates how the lived experience of an individual is deeply connected to larger political structures and social norms which assign value to those experiences. When individuals do not see their realities in data, lack of inclusion in measurement can perpetuate practical social and economic exclusion. And when individuals do not see themselves in data, or there is dissonance between their lived experience and what is reported in available data, this can undermine trust in the data, and, potentially also in institutions associated with that data.

The politics of data, then, is intimately linked to trust in data – or a firm belief that the data are reliable, portray ‘the truth’ or at least identifiable aspects of reality, and do not put respondents’ privacy and security at risk. For some, trust in data may be affected by the end use of such information. In fragile settings, respondents may fear identifying themselves as a specific religion or certain ethnicity group will result in negative consequences for them later down the line. A better understanding of the causes of a lack of trust is critical to strategically planning a path to building and strengthening trust in data. Integrating users and their perspectives into all aspects of the data production cycles, while also ensuring understanding of the methods and intentions, should enhance public trust.

There are important building blocks underlying trust in data that need to be cultivated, towards a data system fit for our future needs. Such building blocks include building strong organizational infrastructures, good governance practices, and multi-stakeholder partnerships; inclusion; following sound and internationally agreed data standards and principles for production, sharing, interoperability, and dissemination of data and statistics, which recognise that data producers and users are increasingly diverse; being transparent about methodology and rationale when developing new approaches that are not yet internationally agreed; and integrating users in the data value chain to meet their needs, increase the value of data and statistics, and ensure meaningful impacts. All of these components can support better use and impact of disaggregated data that will leave no one behind.

Addressing these challenges to find a way forward

Though each distinct in its own way, the four challenges above are all related. Disaggregated data is key to realising the commitment to leave no one behind. For it to be accessible and utilised, it needs to be made open while ensuring privacy and safeguarding trust. The purpose of this event is to discuss each of these issues with a diverse range of stakeholders who offer distinct perspectives and experiences regarding navigating the politics of open data. A topic as complex as this requires a range of viewpoints, including

statisticians, international standard setting agencies such as the United Nations Statistics Division, development data experts, sector professionals such as those working in gender equality and civil society organizations.

This session is designed to start a conversation with a range of stakeholders about the political sensitivities surrounding open data and to answer three main questions:

- What are the major political risks or sensitivities in opening disaggregated data?
- How can these risks be reduced to realise the full value of disaggregated data?
- What are strong examples of ongoing activities, projects, or organizations that are successfully navigating the politics and reducing risks?

An outcome document outlining the main points of discussion and next steps will be shared in the weeks following the HLPF. The document will provide valuable insight for open data champions and advocates who are attempting to navigate the politics of open data. It will share experiences and learning to help in opening data while maintaining trust and privacy; navigating potential sensitivities associated with what new data can reveal and the policy and budget insights that follow; and reducing the challenges and maximizing the opportunities of open, disaggregated datasets. Addressing these challenges can support the global community towards achieving the commitment to leave no one behind.

For follow-up questions about the session please contact Joanna Pradela, Director, Individual Deprivation Measure Team, International Women's Development Agency (IWDA), jpradela@iwda.org.au.