



Inclusive data to leave no one behind – best practices in data disaggregation and use

Claudia Wells, Director of Data Use, Development Initiatives

63rd session of the Commission on the Status of Women

The golden thread of the Sustainable Development Goals (SDGs) agenda is the promise to leave no one behind. Without this commitment, history tells us that only pockets of our society will progress and often, the poorest and most disadvantaged will fall further behind. To realise this promise, we first need data that are timely, comprehensive and disaggregated – to better understand where these people are, their circumstances and how we can capture them in the data we collect. These data then need to be made available to all, from decision-makers and civil society organisations to businesses and citizens so that policies are developed, and governments held to account, in a way that truly reflects the needs of everyone.

Both Development Initiatives (DI) and the Office for National Statistics (ONS) have partnered with the Global Partnership for Sustainable Development Data (GPSDD) and its [global network](#) to reaffirm commitments for improved and strengthened data disaggregation through the Inclusive Data charter. Collectively, the global network will work to improve the quality, quantity, availability and use of inclusive data and share efforts so that progress is accelerated.

What does inclusive data mean?

Inclusive data means ensuring that data are collected for all people, regardless of their location, ethnicity, gender or age. It's about closing the data gaps that inadvertently facilitate discrimination and bias in monitoring, evaluation and decision-making for all of society. And it is especially about ensuring that the use of data is also inclusive.

“Good data makes it intolerable to do nothing” John Pullinger, UK National Statistician

The Inclusive Data Charter is underpinned by five principles:

- Principle 1: All populations must be included in the data
- Principle 2: All data should, wherever possible, be disaggregated in order to accurately describe all populations
- Principle 3: Data should be drawn from all available sources
- Principle 4: Those responsible for the collection of data and production of statistics must be accountable
- Principle 5: Human and technical capacity to collect, analyse, and use disaggregated data must be improved, including through adequate and sustainable financing

The aspiration of this global network is that others will join us in signing up to the charter and share their Inclusive Data Action Plans, so that we are able to share existing learning, demonstrate at a country level how this work is being done – even in challenging contexts – and accelerate the pace of efforts to improve data availability and analysis, enabling the SDGs’ promise to “leave no one behind”.

There are many lessons that can be learnt from the journey to make data more inclusive in both a government and non-government setting.

The P20 Approach – the need for robust comprehensive CRVS and better administrative systems

The P20 Initiative tracks the progress of the poorest 20% of the world’s population from poverty to security and opportunity. It is about ensuring that the Sustainable Development Goals (SDGs) and the data revolution deliver progress for the P20.

People in the P20 face multiple and intersecting deprivations, for example poor nutrition and low levels of education are mutually reinforcing and pass poverty from one generation to the next. People’s identities affect their chances in life with discrimination often reinforcing disadvantage. Existing statistics can help to track national averages, but they often mask disparities at subnational, community and household levels. We need to harness the energy of the data revolution and measure progress by counting people not averages, so that no one – no matter where they live or how old they are and irrespective of their gender, sexual orientation or disabilities – is left behind.

The P20 approach clearly highlights that current Civil Registration and Vital Statistics (CRVS) systems in most countries are not comprehensive enough. Too many people will go through much of their lives being completely invisible on official statistics. Improving administrative data is critical.

Globally, 65% of all births are registered, but among the P20 this is much lower with only one out of every three children (34%) having had their birth registered. With the evidence of identity provided by civil registration, people are better equipped to access education, health services, social protection and employment, to open a bank account and buy or sell assets such as land. CRVS is fundamental to women’s empowerment, increasing independent control over property, inheritance and family relationships. Children who

have been registered are better protected from early marriage, child labour and exploitation. The comprehensive nature of a well-functioning civil registration system, which records every birth and every death, means that no one can be invisible, and policymakers can see the 'universe of need'. If a child's birth is not registered, their death is also unlikely to be recorded, leading to inaccurate estimates of progress on infant or child mortality.

Data from CRVS can be paired with administrative data on health, education and other critical sectors to help plan, deliver and monitor basic services and basic rights.

Why intersectionality between identities matters

Disaggregating data is important for all sectoral analyses but one example of why it is especially important is when evaluating progress in education and understanding where additional emphasis is required. Gender is a critical lens for education as girls are often at risk of not accessing the same educational advantages as boys. But other factors such as income quintile, rural and urban settings, and even birth order are also significant.

For example, disaggregated data showing national average out-of-school rates in Nigeria reveals very different proportions of out-of-school girls and boys, rural and urban children, showing that more out-of-school children are girls, rural based and living among the poorest families, and that identity – in this case ethnicity – results in further exclusion.

Making the most of the data you have, ensure it is used and open

Officials from Uganda Bureau of Statistics (UBOS) confirmed that as of early 2019, only about 5% of published data they have from censuses and various surveys is actually used. Ask yourself this – do you know how much disaggregated data is collected already, would you know which surveys to go to for robust disaggregation? Do you know which of those surveys uses 'best practice' methods to identify at-risk groups?

In the UK, for instance, we have increased the number of disaggregated SDG indicators by making better use of data that are already collected. Data has been made available for [SDGs Goal 16 – Peace and Justice](#). This data highlights the proportion of women held on remand in custody is higher than that for men as women in prison on remand can be held there due to a lack of available appropriate accommodation for them rather than being a threat to the public or committing a particular offence.

Where single years of survey data haven't been robust enough to provide disaggregation, we have used data that combines years. A combined three-year dataset of the Crime Survey for England and Wales was used to provide a large enough sample to identify [women at risk of experiencing partner abuse](#), in line with reporting requirements for SDGs Goal 5 – Gender Equality.

Results showed that young women (under 25), women who have a long-term illness or disability, bisexual women, women who identified with Mixed/Multiple ethnicities, women living in households with an income of less than £10,000 and women living in social

housing all had higher risks of partner abuse. This analysis helps to inform policies and services aimed at ending violence against women and girls in the UK.

It isn't just better use of official data that helps make data more inclusive. The official bulletin on [Domestic Abuse](#) produced by ONS reports data from the Crime Survey and Police Recorded Crime (both official data sources) alongside data from victim support services. The analysis doesn't just look at the prevalence of cases but also highlights how agencies within the criminal justice system and service sector respond to victims and perpetrators of domestic abuse.

In order to address some of the ongoing data needs and gaps ONS has launched a [Centre for Equalities and Inclusion](#). The centre will work with others to make better use of existing data sources and to develop new ones where necessary, taking advantage of new opportunities arising from the data revolution to ensure that the right data are available to address the main social and policy questions about fairness and equity in UK society. Its first release was [an audit](#) of equalities data within government sources.

Within DI, we are also undertaking data audits within countries at subnational level to increase awareness of the data that is available and highlight any gaps. The findings of these audits are being used to advocate for opening up available microdata and increasing investment to fill gaps.

Inclusive Data Use

Inclusive data isn't just about the supply of data, having the capability and confidence to work with data is equally important, especially for women. Data matters, and the need for data-savvy women in leadership has never been greater, yet women are systematically underrepresented in roles related to technology, data, science and engineering.

DI is working within its projects to address the inclusivity of the demand for statistics. Just last month, the Data for Development in Nepal programme, implemented by The Asia Foundation in partnership with DI, hosted Nepal's inaugural Women in Data conference. The [one-day conference](#), which had the theme 'where two superpowers meet', saw an inspiring cast of female speakers, influential panellists, data professionals and aspiring young women join together to discuss the achievements of women who work with data, and how best to further advance female inclusion in this growing specialism. The conference was preceded by a Data Talk from the [Gapminder Foundation](#), and followed by a four-day programme of women in data training sessions. The programme is not just about empowering and building capability in women but also promoting open, accessible and machine-readable formats that allow people to use, reuse and share data without any restrictions.

Recommendation

The challenges to providing inclusive data are huge, but we do not face them alone, and we must not be discouraged by barriers. Instead we must find ways to address them or

get smarter at overcoming them. This is why it is so important that we join together, share our experiences, learn from each other and invest in capability.

My 5 key lessons are:

- Invest in robust and comprehensive CRVS systems
- Consider intersectionality to identify the most vulnerable groups
- Make the best use of the data you have
- Call for data that is collected to be open and accessible
- Invest in the capability of users