



BASIC PRINCIPLES FOR THE **DATA REVOLUTION** FOR SUSTAINABLE DEVELOPMENT

The data revolution will need to be harnessed for sustainable and inclusive development through proactive measures and guided by the following **KEY PRINCIPLES:**

1 **» DATA QUALITY AND INTEGRITY**

Poor quality data can mislead. The entire process of data design, collection, analysis and dissemination needs to be demonstrably of high quality and integrity. Clear standards need to be developed to safeguard quality, drawing on the UN Fundamental Principles of Official Statistics and the work of independent third parties. A robust framework for quality assurance is required, particularly for official data. This includes internal systems as well as periodic audits by professional and independent third parties. Existing tools for improving the quality of statistical data should be used and strengthened, and data should be classified using commonly agreed criteria and quality benchmarks.

2 **» DATA DISAGGREGATION**

No one should be invisible. To the extent possible and with due safeguards for individual privacy and data quality, data should be disaggregated across many dimensions, such as geography, wealth, disability, sex and age. Disaggregated data should be collected on other dimensions based on their relevance to the program, policy or other matter under consideration, for example, ethnicity, migrant status, marital status, HIV status, sexual orientation and gender identity, with due protections for privacy and human rights. Disaggregated data can provide a better comparative picture of what works, and help inform and promote evidence based policy making at every level.

3 **» DATA TIMELINESS**

Data delayed is data denied. Standards should be tightened and technology leveraged to reduce the time between the design of data collection and the publication

of data. The value of data produced can be enhanced by ensuring there is a steady flow of high-quality and timely data from national, international, private big data sources, and digital data generated by people. The data cycle must match the decision cycle.

4 **» DATA TRANSPARENCY AND OPENNESS**

Many publicly-funded datasets, as well as data on public spending and budgets, are not available to other ministries or to the general public. All data on public matters and/or funded by public funds, including those data produced by the private sector, should be made public and “open by default”, with narrow exemptions for genuine security or privacy concerns. It needs to be both technically open (i.e., available in a machine-readable standard format so that it can be retrieved and meaningfully processed by a computer application) and legally open (i.e., explicitly licensed in a way that permits commercial and non-commercial use and re-use without restrictions). The underlying data design and sampling, methods, tools and datasets should be explained and published alongside findings to enable greater scrutiny, understanding and independent analysis.

5 **» DATA USABILITY AND CURATION**

Too often data is presented in ways that cannot be understood by most people. The data architecture should therefore place great emphasis on user-centred design and user friendly interfaces. Communities of “information intermediaries” should be fostered to develop new tools that can translate raw data into information for a broader constituency of non-technical potential users and enable citizens and other data users to provide feedback.

» DATA PROTECTION AND PRIVACY

As more data becomes available in disaggregated forms and data-silos become more integrated, privacy issues are increasingly a concern about what data is collected and how it is used. Further risk arises where collectors of big data do not have sufficient protection from demands from State bodies or interference from hackers. Clear international norms and robust national policy and legal frameworks need to be developed that regulate opt-in and opt-out, data mining, use, re-use for other purpose, transfer and dissemination. They should enable citizens to better understand and control their own data, and protect data producers from demands of governments and attacks by hackers, while still allowing for rich innovation in re-use of data for the public good. Within the agreed privacy constraints, people's rights to freedom of expression using data should be protected. People who correctly provide, collect, curate and analyse data need freedom to operate and protection from recrimination.

» DATA GOVERNANCE AND INDEPENDENCE

Many national statistical offices lack sufficient capacity and funding, and remain vulnerable to political and interest group influence (including by donors). Data quality should be protected and improved by strengthening NSOs, and ensuring they are functionally autonomous, independent of sector ministries and political influence. Their transparency and accountability should be improved, including their direct communication with the public they serve. This can include independent monitoring of the same public services, for example, or monitoring of related indicators such as public satisfaction with services.

» DATA RESOURCES AND CAPACITY

There is a global responsibility to ensure that all countries have an effective national statistical system, capable

of producing high-quality statistics in line with global standards and expectations. This requires investments in human capital, new technology, infrastructure, geospatial data and management systems in both governmental and independent systems, as well as information intermediaries. At the same time, national capacity for data science must be developed to leverage opportunities in big data, to complement high-quality official statistics. Increased domestic resources and international support for developing countries are needed to have the data revolution contribute to sustainable development. Applications of big data for the public good must be developed and scaled up transparently, demonstrating full compliance with applicable laws.

ALL PUBLIC DATA SHOULD BE 'OPEN BY DEFAULT'

» DATA RIGHTS

Human rights cut across many issues related to the data revolution. These rights include but are not limited to the right to be counted, the right to an identity, the right to privacy and to ownership of personal data, the right to due process (for example when data is used as evidence in proceedings, or in administrative decisions), freedom of expression, the right to participation, the right to non-discrimination and equality, and principles of consent. Any legal or regulatory mechanisms, or networks or partnerships, set up to mobilise the data revolution for sustainable development should have the protection of human rights as a core part of their activities, specify who is responsible for upholding those rights, and should support the protection, respect and fulfilment of human rights.