Open data

Introduction

This part of the NSDS Guidelines is concerned with the issues that managers of national statistical systems (NSS) should consider to make statistical agencies more open and transparent and to make statistical information more accessible and useful. It covers concepts and definitions of openness and how these are increasingly used to promote accountability and transparency of government. While the focus of the discussion is on how statistical systems, agencies, and operations can be made more open, the discussion is placed within the context of government-wide initiatives and standards.

This Chapter has benefited from contributions from Open Data Watch and the World Bank.

Open Governments and Open Data

What does the term "open" mean?

National statistical systems are part of central government and the move to promote open statistical data is part of a broader process to make governments generally more transparent, efficient, and democratic and to foster innovation and growth. The International Open Data Charter (1) defines open data as "…digital data that is made available with the technical and legal characteristics necessary for it to be freely used, reused, and redistributed by anyone, anytime, anywhere."

This definition is made operational through six open data principles. The first four describe the legal and technical characteristics of open data:
1. Open by default
2. Timely and comprehensive
3. Accessible and usable
4. Comparable and interoperable

The last two principles describe the purpose and uses made possible by open data:
5. For improved governance and citizen engagement
6. For inclusive development and innovation

Since its 2016 launch, 16 national governments, including those of 10 developing countries, have adopted the International Open Data Charter as a statement of their commitment to open data.This standard of openness can be applied to many of the activities and outputs of governments, academic institutions, and, in some cases, private activities that have been dedicated as public goods.

There are several international initiatives that aim to promote openness and transparency in the public space. The Open Government Partnership (2) has grown from an initial set of 11 countries in 2011 to 75 participating countries now, including 25 developing and middle-income countries. The principles of open government emphasize that openness is essential for making societies more inclusive, just, and sustainable and for promoting economic, social, cultural, civil, and political rights for everyone.

Within this broad movement, open data clearly plays an essential role. As countries develop and as technology advances, both the range of data that governments deal with and the

amounts generated are increasing rapidly.

(1) The Open Data Charter is an international initiative and network that aims to promote and support open data in all its aspects.
(2) Open Government Partnership.

Why Open Data?

Even without a formal government-wide commitment to open data, there are many potential advantages to an NSS adopting formal open data policies and practices, although there are also some important challenges.The aim is to make all kinds of data that a statistical agency produces – aggregate indicators and other statistics, micro-data about individual entities, and geo-spatial data – more open, while still meeting the requirement that data about individuals should remain confidential.

Benefits to NSSsinclude the following (3) .

- Making use of new data sources accessible through other open data initiatives.
- Improving trust in official statistics.
- Improving quality of official statistics.
- Making data sets available for research, analysis, and other social and economic benefits.
- Making data more accessible to users.
- Increasing recognition of the value of official statistics as a public good and as an economic resource

Open data is also a driver of economic growth and job creation (4). Studies show that fast-growing economies often base their success on rich information, which translates into knowledge and more complex and diverse products (5). The McKinsey Global Institute (6) estimates that the US government data helped private firms generate revenues of at least $24 billion annually far in excess of spending on official statistics. They further estimate the potential global economic benefits of open government data may amount to three trillion dollars annually. Another study by a UNECE-led task force estimates that there is a return of US$ 17 for every US$ 1 invested in opening government data on school performance (7).

(3) World Bank, Open Data Challenges and Opportunities for National Statistical Offices.
(4) Open Data Institute, The Value of Open Data.
(5) Hidalgo, Why Information Grows.
(6) Manyika et al., Open Data.
(7) UNECE, Methodologies for Estimating Value of Official Statistics.

Open Data and Official Statistics

The Fundamental Principles of Official Statistics do not explicitly mention open data, but Principle 1 recognizes the obligation of national statistical offices to provide information to all citizens.

Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public

information (8).

National statistical systems have always been concerned with being open and transparent about what they do and the statistics they produce. The value of official statistics is enhanced the more they are used, and making data accessible and useable is an important part of the functions of any statistical agency. In countries where there is already a formal policy or commitment to open government, there is clearly an important role for the national statistical system, and in many cases they will already be actively implementing open data programs and policies. In cases where open government initiatives do not include official statistics, managers of statistical agencies should try to build linkages with and benefit from the broader open data initiative covering all parts of the government. In countries where the Open Data Charter has not yet been adopted or where the national government has yet to develop a formal policy, however, it will still be important for the managers of national statistical systems to identify where and how they can make their agencies and their statistics more open.

This note, therefore, provides guidance on how open data can be incorporated in the preparation and implementation of an NSDS.

(8) United Nations, [Fundamental Principles of Official Statistics](#).

Including Open Data in an NSDs

The responsibilities of national statistical agencies in an open data program include making the data they compile and publish more open, taking proactive steps to ensure their data can be freely used, reused, and redistributed by anyone, anytime, anywhere, subject to any required attribution and maintaining the confidentiality of data about individuals.

It is expected that countries will plan for open data as part of a wider strategic planning process, covering all stages from data production to dissemination and management of statistical assets. In this case, open data concerns would be incorporated into each of the main phases of preparing an NSDS. In general, this will involve three sets of activities: (i) an assessment of the status and strengths and weaknesses; (ii) preparing an implementation plan; and (iii) developing a process for monitoring progress and evaluating outcomes.

Stage 1: Assessment

There are several tools that can be used to assess the extent to which a national statistical system and specific statistical agencies are following an open data approach and to identify where changes and additional actions may be needed.Some of these are designed to assess open data across the whole of government, while others are concerned specifically with the activities of national statistical systems and agencies.

Tools and assessment frameworks that could be used include:

- The Open Data Readiness Assessment (ODRA) (9) , which is an important component of the World Bank's Open Data Toolkit. The ODRA allows an assessment of readiness for open data (as opposed to an assessment of the maturity or effectiveness of an open data program). It is a quick diagnostic and action plan builder, tailored to each country's legislation, institutions, and demand/supply of data. The ODRA includes analysis and recommendations in eight areas: leadership, policy/legal framework, institutions and capacities, data supply, data demand, civic

engagement and capability, funding, and national IT infrastructure. ODRA assessments and the resulting reports are joint products of a small government team and a World Bank team. The methodology can be applied at the national, sub-national, city or agency level. Sectoral adaptations of the methodology have been done for energy and transport.

- The Open Data Barometer (10), developed by the World Wide Web Foundation, an expert assessment, currently available for 92 countries, which provides an evaluation of implementation of open data standards for 15 categories of data, of which national statistics is one. The annual report, available on line, analyses global trends and provides comparative data on countries and regions using an in-depth methodology that combines contextual data, technical assessments, and secondary indicators.

- The Common Assessment Framework (11), also developed by the World Wide Web Foundation, which aims to be of use to countries and civil society organisations with an interest in open data. It is adaptable for use in different environments. It covers the context and environment for open data; the nature and qualities of open data sets; the use of these data sets; and the impact that using open data has in the local environment.

- The Open Data Inventory (ODIN) is an assessment tool, developed by Open Data Watch (12), which focuses on the data compiled and disseminated by national statistical systems through their principal websites. It provides an annual assessment of the extent to which open data principles are practiced. ODIN assessments are available for most developing countries. The assessment covers 20 data categories and 10 main elements under two headings: coverage and openness. The results are tabulated to allow comparisons across different data categories both within a country and between countries.

(9) World Bank, "Readiness Assessment Tool."
(10) World Wide Web Foundation, Open Data Barometer.
(11) Davies, "Towards Common Methods for Assessing Open Data."
(12) Open Data Watch, Open Data Inventory.

Stage 2: Implementing open data programs

Based on the results of the assessment, the next stage is to design and implement specific open data programs across the whole of the national statistical system or for individual statistical agencies. The specific elements of the program will depend on the assessment results, but in line with the NSDS generally, it is recommended that countries should consider three key elements.

First, it is helpful to identify a limited number of actions that can be put into place relatively easily and quickly. Experience from many countries suggests that some or all of the following actions have made an important contribution to making statistical systems more open:

- Advocating towards data producers and other stakeholders on the benefits of opening the data. The human factor is a key element, and tools and processes may not succeed without ownership and real willingness to actually share the data
- Strengthening privacy by design approaches within statistical production processes

- Establishing and strengthening processes for dealing with statistical disclosure control
- Establishing national data archives comprising both metadata and data sets from already completed surveys and censuses (13)
- Establishing and strengthening processes for managing and disseminating metadata, making use of international standards (14)
- Improving interaction and communications with data users, both on-line and through user councils and forums
- Improving data dissemination, especially through statistical websites, for example, by proving for direct downloads of indicators and data series in machine-readable and non-proprietary formats
- Documenting and disseminating information about data policies and practices, providing regular briefings to data users, and maintaining usage statistics to inform user engagement
- Adopting and publishing terms of use that conform to accepted standards for open data, such as the Creative Commons by Attribution (CC-BY) license (15).

Second, an open data program should incorporate a wide range of partnerships, beyond the main stakeholders.It may be possible, for example, to work with software developers to develop applications to help users analyze and use different sets of statistics. However, it will be important to avoid the proliferation of different types of software and to ensure that new tools can be supported and maintained. One area that is often effective for data users is to present more information on maps or infographics. One of the challenges for NSSs is to find the appropriate IT solution. Guidance on this can be found in the World Bank paper, "Technical Assessment of Open Data Platforms for National Statistical Organizations." (16)

Third, an important concern will be to ensure that the sequence of reforms and actions is carefully managed. In many NSDSs the focus of the implementation plan is usually on the first two or three years,even where the formal time frame is five years. It may be helpful, therefore, to develop an open data focused program that uses the same time frame as the NSDS for budget and expenditure planning as well as operational activities.

(13) The [International Household Survey Network (IHSN)](#) has developed tools and guidelines for many aspects of survey management, including data management and the establishment of data archives.
(14) For example, the [Data Documentation Initiative](#) for census and survey data and the Statistical Data and Metadata Exchange for aggregate and time-series statistics
(15) Creative Commons, [About the Licenses](#)
(16) World Bank Group, [Technical Assessment of Open Data Platforms for National Statistical Organisations](#).

Stage 3: Monitoring and sustaining progress

In completing the assessment and preparing an action plan, agencies and NSS will have already identified their short- and longer-term objectives in relation to open data. Therefore, the monitoring process will establish steps to identify changes over time, report on what has been achieved, and identify where further changes in the plans may be needed. Some of the processes and mechanisms that can be used are as follows:

- Monitoring the availability of open data sets (potentially as a share of all data sets available in the NSS)
- Getting regular feedback from data users through regular formal user satisfaction surveys as well as qualitative feedback through user for a and meetings

- Keeping track of usage statistics (e.g. from websites)
- Monitoring the use of micro-data sets
- Monitoring data quality (coverage, timeliness, documentation, etc.)
- Making use of national monitoring processes perhaps supplemented by using assessments tools periodically

Integrating Open Data in the NSDS preparation

Developing an open data approach within the national statistical system will generally require analysis and action plans across many areas of statistical activity within the NSS for each statistical agency and for individual processes and the outputs they generate. Open data actions should be closely linked to other processes, especially the development and implementation of quality assurance frameworks.

It will be important, therefore, to include open data concerns in other components of the overall NSDS preparation process.Where national statistical systems are making use of models such as the Generic Statistical Business Process Model (GSBPM), or the more recent Generic Activity Model for Statistical Organizations (GAMSO), which extends the GSBPM, it will be helpful to incorporate open data actions into these frameworks as well. In this way, the extension of an open data approach to statistical activities should be seen not as a separate activity, but rather as an integral part of designing and implementing statistical activities of all kinds.

Key areas that are likely to feature open data action items in any NSDS will include the following.

- **Managing**–Identify who within the NSDS team will be responsible for ensuring that open data issues are considered and included
- **Committing** –What high-level support is needed to drive the open data agenda and making sure it is delivered
- Budgeting and financing – Ensure that the NSDS budget reflects the actions that are proposed to deliver the open data agenda
- **Advocating** – Build support and ownership for open data throughout the national statistical system and with the main stakeholders
- **Monitoring, evaluation, and reporting** – Open data concerns should be highlighted in the overall monitoring and reporting process

Tools:

The [Advanced Data Planning Toolkit (ADAPT)](#) helps countries plan for meeting the data demands of the SDGs. This tool contains costing modules that plan for open data collection and advocate for more financing. This tool also informs a consultation process for defining a monitoring framework for development and helps chart gaps in financing, data, reporting, and disaggregation. ADAPT was created as a resource for NSOs by PARIS21.
The [Creative Commons Licenses](#) are tools that allow NSOs to grant permission for freely using and reusing official statistics. There are different types of licenses to serve the dissemination objectives of each NSO.

The [Global Open Data Index](#) is an education and engagement tool for citizens to understand the state of open government data in their country. Through a crowd-sourced survey, this tool provides a civil society audit on the openness of government datasets. This tool is

provided by Open Knowledge International.

The [International Household Survey Network (IHSN)](#) has developed tools and guidelines for many aspects of survey management, including data management and the establishment of data archives.

The [Open Data Barometer](#) is a tool to assess the prevalence and impact of open data initiatives. It gathers comparative data to rank countries on readiness, implementation, and impact of open data initiatives within countries. This tool is produced by the World Wide Web Foundation.

The [Open Data Inventory (ODIN)](#) is an assessment tool that scores the coverage and openness of official statistics to identify gaps, promote open data policies, improve access, and encourage dialogue between national statistical offices (NSOs) and data users. Scores allow comparisons across topics and countries. This tool was created by Open Data Watch.

The [Open Data Readiness Assessment (ODRA)](#) is a methodological tool to assess the readiness of a government or agency to implement an Open Data Initiative. It evaluates the commitment of senior leadership to open data, the strength of the legal framework for statistics, readiness for the management of an Open Data Initiative, statistical output and procedures, the demand for open data, the government's multi-dimensional role within an open data ecosystem, funding the supply and demand side of an open data initiative, and the level of infrastructure in place to support an open data initiative. This tool is a component of the World Bank's Open Government Data Toolkit.

The [Open Government Data Toolkit](#) is a set of resources to help governments understand open data as well as plan and implement an open data initiative. This toolkit provides key research and information on starting an open data initiative, exploring available technology options, promoting demand and engagement with users, improving the supply and quality of data, conducting a readiness assessment, and seeking technical assistance and funding. This toolkit is provided by the World Bank.

The [Open Government Guide](#) is a resource for people working to make their governments more transparent, responsive, accountable, and effective. It was developed by the Transparency and Accountability Initiative with the support of the Open Government Partnership.

Good Practices:
AUSTRALIA: [The Corporate Plan 2016-17](#) of Australia is a national statistical strategy that offers a good example of prioritizing collaboration and partnerships and harnessing new data sources.

AUSTRIA: The [Strategy 2020](#) of Austria is a national statistical strategy that discusses improvements to open data and microdata, encourages dialogue between internal and external exports, and seeks to harness new data sources.

CANADA: [The Corporate Business Plan](#) of Canada is a national statistical strategy that seeks to improve accessibility of data through open data. It mentions efforts to increase user outreach and the development of an open data initiative.

CZECH REPUBLIC: The [Czech Republic Open Data Initiative](#) seeks to ensure transparent data infrastructure of the public administration. It emphasizes the need to publish data in open formats.

DENMARK: The Open Data Innovation Strategy (ODIS) of Denmark is an open data policy that seeks to create easier and more uniform access to public data.

GEORGIA: The National Strategy for the Development of Statistics in Georgia (2011-2014) is an NSDS that places significant focus on improving the use of statistics. It emphasizes the need for dissemination policies, open format publication, and dialogue with users.

INTERNATIONAL: The Data Documentation Initiative is an international standard for census and survey data.

KENYA: The Kenya Open Data Initiative shows high-level commitment to open data through the President of Kenya's launch of the government data portal.

MEXICO: The Mexico Open Data Readiness Assessment finds a good state of readiness, identifying the strong coordinating role of the Instituto Nacional de Estadística y Geografía (INEGI) as a strength. The National Digital Strategy of Mexico is an open data policy that highlights open data as a key enabler for coordination, policy implementation, and promoting citizen engagement.

MOLDOVA: The Strategy of National Statistics Development for the period 2008 - 2011 of Moldova is an NSDS that offers a strong example of the assessment, ICT implementation, and focus on online data access.

MONGOLIA: The Program of Official Statistics Development (2006-2010) of Mongolia is an NSDS that discusses the creation of databases to provide official statistics to users through online open data.

NAMIBIA: The Strategic Plan 2012/13 to 2016/17 of Namibia is an NSDS that commits to enhancing online access of official statistics in standardized open formats.

PHILIPPINES: The Open Data Philippines Action Plan 2014 - 2016 describes the introduction of an open government data portal, which will provide data in open formats with metadata and open terms of use.

SERBIA: The Master Plan: Official Statistics of Republic of Serbia – Development and Harmonization 2006 - 2008 is an NSDS that discusses developing the NSO website and improving user access through open data. The Serbia Open Data Readiness Assessment finds a good state of readiness in strong part due to its e-government strategy and open government action plan.

SIERRA LEONE: A National Strategy for the Development of Statistics (NSDS) in Sierra Leone 2008 - 2012 is an NSDS that highlights open data as a critical need of the NSS and covers steps to promote online data access. The Sierra Leone Open Data Readiness Assessment cites the need to strengthen the role of Statistics Sierra Leone (SSL) with increased representation and involvement in open data activities.

The Statistical Data and Metadata Exchange is an international standard for aggregate and time-series statistics.