REPORT

Report of the Expert Group Meeting on the uses of Big Data for Official Statistics: Data Governance and Partnership Models

16 August 2021, 12:00 am to 13:30 pm (Bangkok time)
19 August 2021, 12:00 am to 15:00 pm (Bangkok time)
30 August 2021, 14:30 am to 16:00 pm (Bangkok time)
and
2 September 2021, 12:00 am to 15:00 pm (Bangkok time)

ESCAP 75
MOVING FORWARD TOGETHER

About the report

During 16 August – 2 September 2021, ESCAP Statistics Division organized the Expert Group Meeting on Big Data for Official Statistics: Data Governance and Partnership Models. The event was aimed at the following – exploring interest and need for regional guidance to NSOs on big data governance and partnerships; exploring institutional and legal foundations for successful big data integration into official statistics; and country experience sharing. The EGM focused on two topics: Big Data Governance and Big Data Partnership models and was comprised of four sessions: two public Stats Cafes and two closed expert discussions - one dedicated to each topic.

The two Stats Cafes on Data Governance and Big Data Partnership Models were attended by over 150 online participants each. The presentations and recordings of the Stats Cafes are publicly available and fed into the succeeding expert discussions. The two expert discussions were attended by 30 participants from the national statistical offices, development partners and private sector.

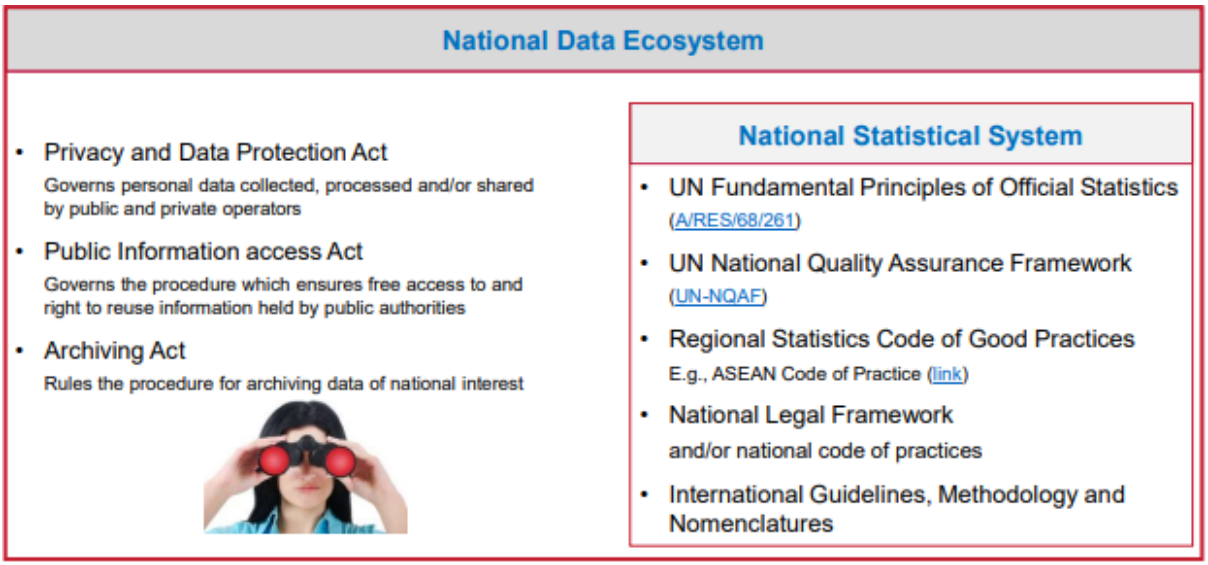
This report highlights the main points presented and discussed during the four sessions of the Expert Group Meeting

Table of Contents

I. NATIONAL STATISTICAL REGULATORY FRAMEWORK

The National Data Ecosystem goes beyond the National Statistical System (as seen in Figure 1) and comprises privacy and data protection and access to public sector information and archiving acts. These acts guide use and reuse of data at the national level. A national legal framework lays out the rules and data sources for the production of official statistics. Furthermore, regional and global principles, guidelines and frameworks provide additional guidance to national statistical offices.

Figure 1. National Data Ecosystem



Source: Gabriel Gamez, Principles, Codes and Normative Frameworks, Stats Café presentation on 16 August 2021

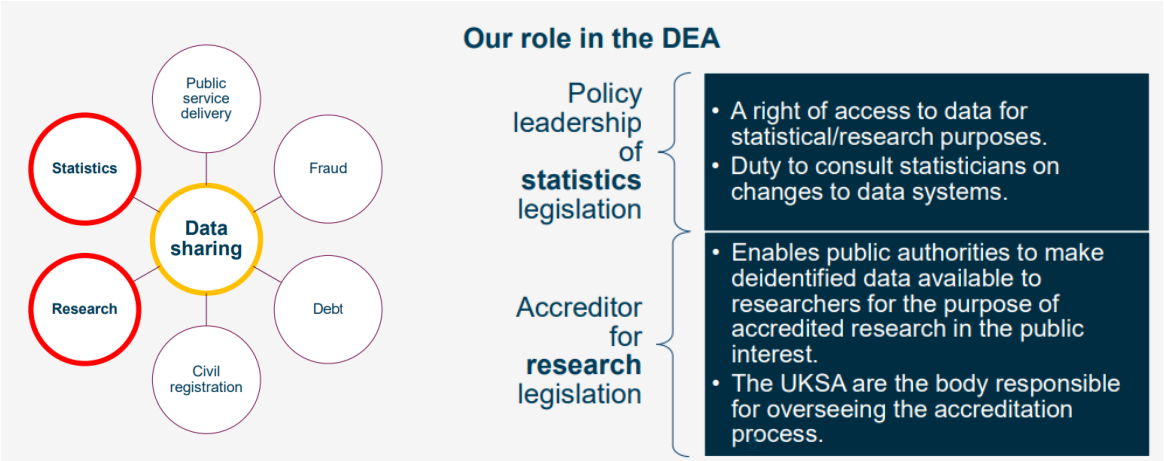
In many countries, the **statistical legislation is outdated and does not support the production of official statistics with alternative data sources**, despite strong pressure to produce more granular and timely data. In countries with capacity and resource constraints, some NSOs are pressured to use alternative sources of data to produce official statistics. In others, there is a pronounced lack of political will from the leaders to use secondary data sources in decision making. Hence, countries have varying national regulatory framework and experience with integrating alternative sources of data into official statistics.

Few countries manage the overall access to non-traditional data sources through national legislation. Others try to regulate access and use of individual data sources separately. However, the use of big data for public good could also be considered beyond official statistics; it could be considered for improving public services, research and other areas. Ideally, governments and NSOs should work towards a one consolidated data strategy that would guide data collection, use and reuse across both the government and with non-government actors.

Example 1. For example, the United Kingdom's Digital Economy Act (2017) regulates, among others, access to any data source for purposes such as statistics, research, public service delivery, civil registration, fraud identification and debt. (Figure 2). Access to the privacy sector data was also possible through this legislation. Furthermore, the UK's National Data Strategy commitments to driving the use of the Digital Economy Act powers a way to removing barriers to data sharing. Following this strategy, Office for National Statistics (ONS) is leading on the delivery of the Integrated Data Services through cross-government collaboration. ONS took the lead as it is highly trusted by the public and the integrated data approach is a solution to lack of data sharing across the UK government departments.

Example 2. The government of Indonesia has also developed a national data strategy but did not involve the private sector. BPS Statistics Indonesia has developed a data hub combining different data sources. It will address issues related to data quality, lack of standardization, multitudes of platforms and others. In the last four years, however, it has embraced public-private partnerships and is actively exploring and integrating big data into official statistics.

Figure 2. Our role in the DEA



Source: Simon Whitworth, the ONS' role in the Digital Economy Act, Stats Café presentation on 16 August 2021

Confidentiality and personal data protection were identified as the main regulatory constraints to using big data for official statistics (Figure 3). The regulatory framework on personal data protection varies across the region. Countries like Indonesia are approving the national Privacy Act in line with the General Data Protection Regulation (GDPR) provisions. Others follow the statistical legislation or the fundamental principles of official statistics, that guide the preservation of personal data confidentiality.

However, when it comes to **data sharing across different government entities, the current individual ministerial legislation may impede data sharing because of privacy concerns**. This is also true in multiple countries, when attempts are made to access data through the regulator. There are cases where the regulator collects data from the private sector but cannot then share it with NSOs or other public institutions due to legal constraints.

Figure 3. Opinions received during the Expert Discussion on Big Data Governance, 19 August 2021

Some countries in the region are fine-tuning their statistical acts and need support in identifying best practices of integrating big data into the production of official statistics. One of the questions **is how the national statistics offices can ensure that the statistical act has an upper hand over institutional legislation protecting data sharing**.

Example 3. In Indonesia, the Privacy Act is currently being revised in the Parliament and allows the use of individual data for statistical purposes with prior consent. However, the current Statistics Law prohibits individual data sharing, and mandates that administrative data is shared at aggregated data. For research purposes, where micro data is needed, depersonalized data from surveys, census and administrative sources are shared through a Letter of Agreement for data use. When it comes to the use of mobile phone data, a Non-Disclosure Agreement (NDA) is signed with the mobile network operator.

Example 4. In the UK, the Center for Applied Data Ethics provides independent expert ethical advice through self-assessment tools and guidance, and provides an ethics user support service for the research and statistics community at the research design phase.

Continuous compliance with the ethics principles and legislation is of major importance to maintain public trust of the statistical office and the statistical system when big data is used. When conducting pilots and data integration, NSOs should continuously show that they are working in the spirit of the legislation and are abiding to the principles of privacy and ethics. The public should understand how their data are used. However, it was noted that ensuring transparency and accessibility of data requires significant time and resources.

II. SCALING UP BIG DATA PROJECTS

The **transition from big data experiments to big data use for regular statistical production should occur following the piloting phase** to maximize the degree of innovation of the respective application. The NSOs should start by testing many options in a "fail-early" agile way and then build a use case and business case for pursuing the access and use of certain data for a specific application. There are several benefits to such an approach. These include having a small team of methodologists to carry out the pilot, low costs and short timing needed to assess the viability and needs for further data integration, when assessing the transition requirements in terms of data, infrastructure and resources.

While in most cases alternative sources **complement traditional sources of data**, there are cases where they can **replace traditional production of certain statistics**. These cases include using Mobile Positioning Data as a replacement for household survey data to assess domestic tourism in Indonesia, at a third of the household survey cost and with greater granularity and periodicity.

Demand for timely data to inform policy and address public issues can facilitate or speed up big data adoption and integration. For example, in Mexico, which is experiencing fast urbanization, but the census is conducted every 10 years, National Institute of Statistics and Geography (INEGI) is monitoring city expansion through satellite imagery. Similarly, in Australia, satellite data is used for dwelling assessments, saving money and time. Therefore, building a **business case for data use to solve a pressing issue could facilitate faster access to data along with important cost savings**.

III. BIG DATA PARTNERSHIP MODELS

To use alternative sources of data for official statistics, national statistical offices should secure access to those data first. These data can come from both within the government and non-government institutions in the private sector. New models which move beyond the traditional philanthropic initiatives are emerging, as highlighted in Figure 4.

Figure 4. Emerging models for NSOs access to MNOs data

Source: Martina Barbero, Stats Café presentation on 30 August 2021

COMMERCIAL MODEL. NSOs can access private sector data through commercial models, by paying companies full or preferential price for accessing data or services.

Example 5. In the absence of regulatory guidelines on private sector data access, BPS Statistics Indonesia entered a commercial agreement with Telkomsel, an Indonesian wireless network provider and signed a Memorandum of Understanding. The mobile network operator made significant investments in the infrastructure to store signaling data, which is of no business utility to the company. These costs

RECIPROCITY MODEL. Another model is the reciprocity model, where the NSOs identify non-monetary incentives for the private sector, such as producing aggregated data that would allow the company to calculate its market share.

GOVERNMENT LEGISLATION is an approach taken by several countries in Europe. This is an example of UK's Digital Economy Act and the European Data Act.

INTERMEDIATION BY THE REGULATOR. Another approach discussed was the intermediation by the regulator. However, several NSOs pointed to a challenging