

# *Data Stewardship and the secure sharing of data*

## International Seminar on Official Statistics

2-3 November 2022, Seoul, Republic of Korea

### Concept Note

The UN Statistics Division and Statistics Korea (National Statistical Office of the Republic of Korea) have co-organized every year an international seminar regarding important issues on the global statistical agenda since 2010. The seminar has covered a variety of social and economic issues from the perspective of statistics. For example, we have discussed the role of data and statistics in understanding open data and data disaggregation for reaching Sustainable Development Goals; and using big data for official statistics. As we recover from the COVID-19 pandemic, the significance of timely data using new data sources and data linkage has increased and it has provided us with new opportunities to the international statistical community for discussing the data stewardship and the data security issues.

In 2014, the UN Statistical Commission established the [UN Committee of Experts on Big Data and Data Science for Official Statistics](#) to explore the benefits and challenges of using new data sources and new technologies for the community of official statistics. How could we make sure that the national statistical systems would remain relevant in a fast-moving data landscape? Among others, this Committee of Experts looked into the use of satellite data for agriculture statistics and landcover mapping, the use of mobile phone data for tourism, population and transport statistics, and the use of retail scanner data for estimating the Consumer Price Index. Getting access to private sector data is a recurring problem from the beginning because of security and privacy concerns.

#### **Data Stewardship**

In 2021, the UN Statistical Commission created a UN working group on data stewardship to promote sharing of data, whether public-public or public-private, allowing the regular use of relevant data for the generation of official statistics, while preserving privacy and confidentiality; and to improve the national data ecosystems through better methods and access inspired by a data stewardship approach. The program of work of this group covers – among others – data governance, equity and inclusion; the development of a conceptual framework for data stewardship; and increased attention to issues related to protection of privacy of data while ensuring its full utilization.

According to views developed in New Zealand, there are 7 key elements required for effective data stewardship:

- Strategy and culture – a strategy that provides a shared vision and clear direction, and a data culture that enables strategy implementation and sustains good data stewardship practice.
- Rules and settings – legislation, policies, principles, and sanctions providing boundaries and guiding how the data system should operate.
- Roles, responsibilities, accountabilities – governance structures, role definitions and expectations, and leadership.
- Data capability and quality – tools, processes, designs, metadata structures, and platforms for managing, storing, describing, and sharing data.
- People capability and literacy – skills, knowledge, and services for accessing, managing, analyzing, and communicating data and insights.
- Influence and advocacy – effective relationships and networks to endorse, promote, and support good data practice.
- Monitoring and assurance – assessing environmental trends and developments, measuring stewardship performance, and adapting the stewardship toolkit to respond to changing circumstances or new information.

In 2021, Statistics Korea began promoting the establishment of the “Statistical Data Hub Platform (formerly referred to as the K-Statistics System)”, a public big data system leveraging cutting-edge privacy-preserving techniques and ICT to enable the safe linkage and use of scattered governmental data. The more data we collect and link together, the more our economies of scale and scope can be understood. Accordingly, Statistics Korea aims to maximize the potential value of data by facilitating the linkage between governmental data through statistical registers on population, households, and establishments. The Statistics Data Hub Platform of Statistics Korea can be considered as a good national practice of Data Stewardship. The current situation of the Platform will be discussed during the seminar.

### **Secure sharing of data**

Official statistics are a trusted source of information for governments around the world to make informed and data-driven decisions. As such, data are collected from a range of sources, such as household and business surveys, population, economic or agricultural censuses, a variety of administrative records or even private sector data. Those data sources are the inputs for the compilation of statistics and indicators on the economy, the environment and the society. In many ways, official statistics offer a snapshot of a country's development and rate of progress.

Naturally, the more fine-grained the level of input data, the more nuanced the official statistics can be. However, the collection, processing, and dissemination of often sensitive

data should only be done, if the privacy of persons and businesses is protected. So, how can we use technology to mitigate privacy risks and give provable privacy guarantees throughout the collection, processing, analysis and distribution lifecycle of potentially sensitive information?

Protecting data from unauthorized access, processing or distribution is the simple goal of privacy-preserving statistics. With such broad applicability to the day-to-day processes of official statistics, it is in the interest of every official statistics office to have an adequate level of understanding of privacy-preserving statistics. National statistical offices that actively promote and utilize privacy-preserving statistics have the opportunity to build greater trust with its citizens and hence unlock new opportunities associated with more accurate and complete data collection. Statistics Korea has promoted the development of data privacy-preserving techniques such as homomorphic encryption, differential privacy and synthetic data through the National Research and Development Project, which will be leveraged to construct the “Statistical Data Hub Platform” between 2021 and 2024, in cooperation with the Ministry of Science and ICT. There will be a special session organized by Statistics Korea during the seminar for having further discussion regarding the data security issues covered by national statistical authorities.

This seminar will also cover the topic of privacy-enhancing technologies (PETs), designed to safely process and share sensitive data.