United Nations

E/CN.3/2021/17

Economic and Social Council

Distr.: General
21 December 2020

Original: English

Statistical Commission
Fifty-second session
1–3 and 5 March 2021
Item 3 (m) of the provisional agenda*
Items for discussion and decision: data and metadata
presentation and dissemination

Data and metadata presentation and dissemination

Report of the Secretary-General

Summary

   The present report, which was prepared in accordance with Economic and Social Council decision 2020/211 and past practices, provides information about the activities of the Statistics Division of the Department of Economic and Social Affairs of the Secretariat on the presentation and dissemination of data and metadata in collaboration with other United Nations system entities and other international and regional organizations. It contains an overview of the arrangements, technical infrastructure and skills required for a holistic and integrated approach to the presentation and dissemination of statistical data, metadata and other related data to different user groups. The report provides a vision and way forward for the creation of a global United Nations data hub based on modern interoperability standards and a federated data governance architecture. It is also aimed at contributing statistical expertise and standards to the overall Data Strategy of the Secretary-General for Action by Everyone, Everywhere. The points for decision by the Commission include: (a) expressing its views and endorsing the vision and way forward for the implementation of the new UNdata portal; and (b) endorsing the establishment of an inter-agency and expert mechanism for the full development and implementation of a new United Nations data hub and the regular review and update of the data contained in UNdata and its proposed terms of reference.

---

* E/CN.3/2021/1.

20-17544 (E)   311220

I. Overview of the work of the Statistics Division on the presentation and dissemination of data and metadata

1. The Statistics Division of the Department of Economic and Social Affairs of the Secretariat presents and disseminates internationally comparable data and metadata on countries, areas and regions of the world through different products, such as statistical yearbooks and statistical databases. In addition, statistical data is presented in such analytical publications as *The Sustainable Development Goals Report* or *The World's Women*, which provide a written analysis of data, supported by the visual presentation of that data in the form of graphs and maps. Those analytical publications are often complemented by dedicated web and mobile applications that present the same and additional information in interactive webpages. In addition, the Division's statistical databases may be embedded in user portals that contain links to visualizations of data and additional analytical products.[1]

2. In addition to compiling internationally comparable data and metadata for multi-domain products, the Division compiles data and metadata in specific statistical domains, such as trade or demographic statistics directly from countries and areas, often in collaboration with other international or regional organizations so as to avoid any duplication of the reporting burden on countries. In some cases, those data compilations are the only source with global coverage on the specific topic in question.

3. The production of statistical outputs covering data and metadata across different statistical domains, such as the United Nations *Statistical Yearbook*, the *World Statistics Pocketbook*, the *Monthly Bulletin of Statistics* and their respective statistical databases, such as UNdata[2] and the Global SDG Indicators Database, rely on the cooperation and collaboration with other United Nations system entities and other international and regional organizations, which compile and submit the data and metadata from countries in line with their respective mandates. In these cases, the Division redisseminates or reuses data and metadata that has already been published by the respective organizations, adding value by bringing those varied data and metadata together and making it accessible to a wider audience in a standardized format. However, there are cases where the Division is the sole provider of the data and metadata in database format, as some entities do not maintain their own database infrastructure for dissemination and provide, for example, only individual tables for download.[3] In addition, the Division acts as a global and trusted steward of internationally comparable data, assuring users of the quality of the data and metadata provided.

4. The statistical products and data and metadata compilations of the Division are authorized by Member States and serve specific purposes, be it to support the compilation of internationally comparable data and metadata in a specific statistical domain or to monitor and report on economic, social or environmental phenomena supporting the various United Nations areas of work on the global, regional and national levels.

---

[1] See, for example, https://comtrade.un.org.

[2] The UNdata portal is available at http://data.un.org.

[3] See, for example, *World Population Prospects*, available at http://data.un.org/Search.aspx?q=population+datamart[PopDiv] and https://population.un.org/wpp/Download/Standard/Population.

II. Arrangements, technical infrastructure and skills

5.   The compilation of data and metadata in specific statistical domains is based on regular data requests to countries, often in collaboration with international or regional organizations, while the compilation of data and metadata for multi-domain products and databases such as UNdata and the Global SDG Indicators Database relies on the submission of data and metadata by other United Nations System entities and other international and regional organizations, which compile the data and metadata from countries. Data compilation and processing at the Division is facilitated by standardized data templates and data requests, accompanied by detailed instructions on modalities and formats for data submission.[4] Nevertheless, in many cases, data and metadata must be pre-processed before they can be loaded to the Division's databases and, often, data must be manually obtained from the data reporter's respective websites. In many instances, successful efforts have been undertaken to automate data compilation and to use the Statistical Data and Metadata Exchange (SDMX) standard for automated data exchange. However, wider use of automated data exchange or other advanced techniques for data-sharing and data integration are hampered by different levels of technical capacities and skills and the lack of use and availability of common standards.

6.   The Division has developed specialized databases and acquired the technical infrastructure, including server infrastructure and software licences, to support the production and dissemination of its statistical outputs in both print (pdf) and online form. Rapid advancements in information technology, software and tools provide opportunities for the Division to upgrade its capacity for data processing, presentation and dissemination and to offer additional products. At the same time, those advancements and the use of innovative technologies have created an increasingly complex technical environment, which makes maintaining a large number of separate statistical products and processes more difficult.

7.   The maintenance and further development of statistical databases and the production of statistical outputs of the Division is supported by teams consisting of statisticians, information technology (IT) professionals and statistical assistants. Statisticians and statistical assistants have been increasingly acquiring advanced IT skills, while IT professionals have been increasingly involved in statistical analysis tasks, thus creating a pool of staff that fit the profile of data engineers and data scientists. That pool has allowed the Division to maintain and further develop its databases and products, despite an increase in mandates relying on the same resources, and to embrace some of the new technological developments, such as the use of application programming interfaces and automated data and metadata exchange systems, the use of geographic information systems analysis and visualization tools and the use of modern web technologies to provide access to, and facilitate the use of, statistical outputs, despite the Division's very limited resources.

III. A new UNdata portal

8.   The UNdata portal is the central access point for the data and metadata assets of the Division. Data and metadata compiled by other United Nations system entities and international and regional organizations and shared with the Division are also made available through the portal. The UNdata portal draws large numbers of users, and its usage has been increasing over time.

---

[4] An overview of data requests sent regularly by the Statistics Division to countries is available at https://unstats.un.org/home/questionnaire.

9. At its launch in 2005 as part of a project entitled "Statistics as a Public Good", UNdata was highly innovative, establishing a data warehouse that brought together statistical data and metadata from across the United Nations system, making them easily available to users, who could conduct analytical queries from a central location. Since that time, however, new approaches and technologies for data integration, presentation and dissemination, including federated architectures and data virtualization, have become widely available. Those new approaches and technologies allow users to retrieve data from multiple sources without imposing a single data model or requiring data to be physically stored in a central location. Currently, the responsibility for maintaining the various data marts of UNdata rests almost solely with the Division and relies on manual and semi-automated extract, transform and load processes. The inclusion of new datasets is burdensome, as it often requires both the transformation of the original data and the modification of the UNdata schema. From the user perspective, major limitations of the current UNdata platform include issues relating to data and metadata availability, ease of access and data and metadata presentation and visualization. The current UNdata platform presents difficulties in meeting the requirements of users and the advanced needs of the United Nations statistical system with regard to a common platform for data and metadata presentation and dissemination.

10. In the Data Strategy of the Secretary-General for Action by Everyone, Everywhere, a vision is set out for a data ecosystem that maximizes the value of United Nations data assets for the Organization and the stakeholders it serves, including high-level objectives, principles, core workstreams and concrete system-wide data initiatives.[5] The proposals contained in the present report with regard to modernizing UNdata in terms of how it collects, manages, uses and shares data are aligned with the Organization's overarching Data Strategy and would constitute a major contribution from the statistical community to the vision of the Secretary-General.

11. As a contribution to the Data Strategy, the Committee of Chief Statisticians of the United Nations System developed a system-wide road map for innovating United Nations data and statistics (CEB/2020/1/Add.1). The road map is aimed at strengthening the United Nations system as a primary provider of global data and statistics by transforming the data.un.org portal into a single point of reference for the system. One of the action items called for in the road map is the updating of the United Nations data portal so that data from system entities can be linked and presented under a unique United Nations data brand. The road map contains the following proposals for realizing that action item: (a) explore technology to establish an online United Nations data portal, where data identified by United Nations system entities as "United Nations data" can be automatically linked and presented; (b) rebrand United Nations data; (c) update the United Nations data portal with a new scope; (d) define common data and metadata standards; (e) finalize the automatic upload of data; (f) adopt a common open data licence for all resources available through data.un.org; and (g) develop interactive data exploration and visualization tools. Moreover, under the road map, the Statistics Division is called upon to lead in the design, development and maintenance of the UNdata portal, in coordination with United Nations system entities.

12. To implement that vision, it will be necessary to explore new data integration approaches, including, for example, a federated architecture focused on the

---

[5] See www.un.org/en/content/datastrategy/index.shtml.

coordinated sharing and interchange of information among multiple data hubs.[6] This would include developing a modern governance framework and the necessary data management infrastructure to position UNdata as a global network of autonomous, authoritative data hubs that provides visibility to all participating national, regional and international organizations, while allowing users to access data services and analytic tools and enabling both producers and users of data to interact more efficiently with each other in mutually rewarding ways.

13.   It is envisaged that a new United Nations data hub might have the following requirements: (a) to serve as the main entry point to authoritative statistical data, and the corresponding metadata, from the United Nations system and other participating international and regional organizations; (b) to provide a modern architecture and tools for the integration,[7] presentation, analysis, visualization, communication, use and reuse of data and metadata; (c) to ensure the common use of international statistical and geospatial data and metadata standards, concepts and classifications, thus facilitating interoperability, traceability, integration and usability; (d) to implement a data governance framework in which all United Nations system entities and other participating international and regional organizations and data providers share responsibility, taking into account their respective mandates and capabilities; (e) to enable strong engagement between members of the United Nations statistical system and all its user constituencies; and (f) to ensure feasibility and sustainability in the realization of the vision for the United Nations data portal as outlined in the system-wide data and statistics road map.

14.   The feasibility and sustainability of the vison for a new UNdata portal will depend on the mobilization of the required additional resources across the United Nations system, which may, to a large degree, depend on the commitment of data users within the system to treating the statistical data and metadata provided through UNdata as an indispensable and strategic asset, as is the case in other data-driven organizations. As envisioned in the Data Strategy, better approaches to data will deliver better outcomes for everyone: stronger decision-making and policy advice; greater access to and sharing of data; improved data governance and collaboration; more robust data protection and privacy; enhanced efficiency across United Nations system operations; greater transparency and accountability; and better services for people and planet. It is expected that the realization of that vision will contribute statistical expertise and standards to the overall Data Strategy.

IV.   The way forward

15.   The discussion and vision outlined herein provide an introduction to the challenge of developing a new UNdata portal. Further consultations with all stakeholders are required to fully develop and realize that vision. It is therefore suggested that a dedicated inter-agency and expert mechanism be established to enable the participation of relevant United Nations system organizations and other relevant organizations, as well as experts from national statistical systems and other stakeholders, in the full development and implementation of the vision outlined herein and in the regular review and update of the data contained in UNdata.

––––––––––––––

[6] In a federated architecture, each hub maintains ownership over specific data assets and autonomously decides what data to share and how to share them. Moreover, data assets remain at the source, with references to them being passed to other participating hubs to ensure findability and traceability.

[7] Three commonly mentioned modern approaches to data integration are: data lakes; data virtualization or federation; and data hubs.

V. Action to be taken by the Statistical Commission

16.   **The Commission is invited:**

(a)   **To express its views and endorse the vision and way forward for the implementation of the new UNdata portal;**

(b)   **To endorse the establishment of an inter-agency and expert mechanism for the full development and implementation of a new United Nations data hub and the regular review and update of its data and metadata, and to approve its proposed terms of reference.**

Annex

Terms of reference for an inter-agency and expert mechanism for United Nations data

I. Background and mandate

1.   The Inter-Agency and Expert Mechanism for United Nations Data will meet the requirements of the Data Strategy of the Secretary-General for Action by Everyone, Everywhere regarding the presentation and dissemination of statistical data and metadata of the United Nations system and beyond, on a new or upgraded UNdata portal.

II. Objectives and tasks

2.   The Inter-Agency and Expert Mechanism for United Nations data will:

(a)   Review the vision for a new or upgraded UNdata portal and further amend and elaborate that vision, as needed, on the basis of a detailed assess ment;

(b)   Decide on the strategy and steps needed for the implementation of a new United Nations data hub;

(c)   Work on mobilizing the required resources, including with regard to advocacy;

(d)   Request and encourage all data providers to work on the realization of the vision and strategy for a new or upgraded UNdata portal and take responsibility for the availability of the providers' data and metadata on or through UNdata;

(e)   Ensure the regular submission of data and metadata from all entities to the UNdata portal.

III. Membership

3.   Membership is open to all members of the Committee of the Chief Statisticians of the United Nations System, all members of the Committee for the Coordination of Statistical Activities, representatives of national statistical systems and other stakeholders. A representative of the Executive Office of the Secretary-General will be invited as an observer.

IV. Governance and organization of work

4.   The Statistics Division of the Department of Economic and Social Affairs of the Secretariat and one member of the Inter-Agency and Expert Mechanism for United Nations Data, upon election, will be Co-Chairs of the Mechanism. The Mechanism will report biennially to the Statistical Commission.

———————————