Economic and Social Council

Distr.: General
15 December 2023

Original: English

Statistical Commission
Fifty-fifth session
New York, 27 February–1 March 2024
Item 5 (j) of the provisional agenda*
Items for information: data stewardship

Report of the Working Group on Data Stewardship

Note by the Secretary-General

In accordance with Economic and Social Council decision 2023/325 and past practices, the Secretary-General has the honour to transmit the report of the Working Group on Data Stewardship, which is submitted to the Statistical Commission for information.

---

* E/CN.3/2024/1.

Please recycle

Report of the Working Group on Data Stewardship

I. Introduction

1.   At its fifty-first session, the Statistical Commission, in its decision 51/120 (2020), established the work on data stewardship under the High-level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development and asked it to report back at its next session with a stocktaking exercise and specific recommendations on the way forward. As a result of the High-level Group's work, the Commission created the Working Group on Data Stewardship the following year through its decision 52/103 in order to develop further guidance on approaches to data stewardship for national statistical offices. Since then, the Working Group has formed five workstreams to conduct its work, covering in each: data governance (workstream 1); equity and inclusion (workstream 2); sharing and collaboration (workstream 3); links with the city data agenda (workstream 4); and development of an overall conceptual framework (workstream 5). In 2021, it absorbed the outstanding work of the Working Group on Open Data. The Working Group reports to the Commission on an annual basis.

2.   The Working Group is composed of representatives from national statistical offices and stakeholders from other data communities, including international and regional organizations and civil society organizations. The current co-Chairs are the Director General of the Departamento Administrativo Nacional de Estadística (Colombia), Beatriz Piedad Urdinola Contreras, and the President of Statistics Poland, Dominik Rozkrut. The Statistics Division of the Department of Economic and Social Affairs of the Secretariat serves as secretariat for the Working Group.

3.   During 2023, the leads of the workstreams met four times and the Working Group met once. The co-Chairs and other lead members of the Working Group helped to organize and participated in various events during the year. Some examples were a side event, entitled "Data stewardship: from buzz word to a viable governance framework for data and statistics", at the fifty-fourth session of the Commission, a session on the parallel programme, entitled "Data stewardship – comparing country developments", at the recent United Nations World Data Forum, held in Hangzhou, China, a webinar called "Data stewardship in Africa: the role of national statistical offices", organized with the Economic Commission for Africa, and, lastly, a side event organized within the framework of Economic Commission for Latin America and the Caribbean's Statistical Conference of the Americas.[1]

II. Global consultation questionnaire implementation and headline results

4.   The global consultation questionnaire was developed during 2022 and reviewed by the Commission during its last session. The questionnaire was approved by the Commission, with the majority of the 21 written statements received on the report explicitly supporting the questionnaire as it had been presented. However, the Working Group took into account the feedback from one written statement and added an open question on the opportunities and challenges of the data stewardship role in the wider data ecosystem, given that all other questions were either explicitly focused on the coordination role of the national statistical office in the national statistical system

---

[1] Further information on the Working Group's terms of reference, its membership, past events and its written reports are available at https://unstats.un.org/unsd/statcom/groups/data_stewardship/.

or provided a set of functions[2] and a list of perceived successful implementation measures[3] commonly associated with data stewardship but still centred in the national statistical office's experience with the national statistical system. The finalized questionnaire contained eight questions – six mainly quantitative and two qualitative questions – described briefly as follows:[4]

(a) Status of the national statistical office as the coordinator of official statistics in the national statistical system;

(b) Coverage of data stewardship functions in Government, including space to specify those responsible;

(c) Potential role of national statistical office as coordinator of all data sources in the national statistical system, including space to specify reason for answer;

(d) Change in national statistical office responsibilities relating to data stewardship functions in past two years;

(e) Definition and measurement of successful data stewardship implementation;

(f) Current working links to subnational and city levels;

(g) Opportunities and challenges for the national statistical office as data stewards in the national statistical system (open question);

(h) Opportunities and challenges for the national statistical office as data stewards outside the national statistical system (open question).

5. The questionnaire was circulated among national statistical offices on 17 July 2023 with a deadline to complete it of 6 September 2023. A final extension was given until 30 September. The questionnaire was made available in all six official United Nations languages, given that it was translated as part of the Working Group's report to the Commission at its last session. This assisted in the collection of 113 responses from national statistical offices, of which 33 were collected through the non-English versions of the questionnaire. This meant that the global consultation received an overall response rate of 58 per cent from a potential 194 national statistical offices representing Member States and observers. The regional response rate across 22 subregions defined by the Division's geographical regions[5] showed that 16 subregions had a response rate of 50 per cent or better. For the analysis of the six quantitative questions, regional differences will be compared using a set of five regions on the basis of the Sustainable Development Goal grouping.[6] This regional grouping has the

––––––––––––––––––

[2] The list of functions included making data and statistics readily findable and accessible, supporting data-related work across Government, promoting and monitoring good practices for use and reuse, improving governance policies and management practices, building and strengthening partnership and collaboration with the national statistical system/civil society, ensuring privacy and confidentiality, ensuring quality and interoperability, overseeing the security of data systems and data exchange, promoting and leading technological and statistical standards and fostering user engagement.

[3] The list of successful implementation measures included active engagement between producers and users, increased availability and timeliness, increased citizen engagement in surveys and censuses, increased data literacy, increased data-sharing and reuse, increased resources allocated, increased trust, increased use of statistics by media, low incidence or impact of data privacy breaches and more data-driven and statistics-driven policymaking.

[4] See annex I to E/CN.3/2023/20 for exact wording of the questions and available responses, except for question (h), which was added later.

[5] See https://unstats.un.org/unsd/methodology/m49/ for more details.

[6] Sustainable Development Goal regions in the analysis are defined as North Africa and Western Asia, North America and Europe, sub-Saharan Africa, Latin America and the Caribbean and Asia and the Pacific. See https://unstats.un.org/sdgs/indicators/regional-groups/ for more details.

advantage of closely reflecting the regions covered by the system of United Nations economic and social commissions, where significant work on data stewardship is also taking place. The results from the questionnaire in the present report will look mainly at the global highlights, while regional and more detailed results will be reported in a separate background document accompanying the report.

6.   With respect to the first question of the national statistical office coordinator role for official statistics in the national statistical system, some 80 per cent of national statistical offices reported being mandated to coordinate the production and integration of official statistics into the national statistical system. As to whether they believed they should coordinate all data sources within the national statistical system, a very similar percentage agreed. The reasons for supporting the coordination of all data sources under the national statistical system related to both technical and institutional benefits. The technical benefits often related to data quality issues such as managing accessibility, timeliness and coherence of data flows, while the national statistical offices presented themselves as the most appropriate institution within Government with the capacity and expertise to perform that role. For the approximately 12 per cent that disagreed with the national statistical office taking a coordinator role for all data sources, the reasons given were perceived boundaries between data and official statistics, capacity limitations to cover such a diverse range of data and other constraints such as the decentralized nature of their statistical system. The institutional aspects related to the improvement in data governance policies and data management practices. Overall, there was a consistent opinion that the coordination of data sources was an extension of the national statistical office's role as coordinator of official statistics.

7.   The majority of national statistical offices indicated that authorities or bodies within their countries were already responsible for performing the perceived list of data stewardship functions. National statistical offices reported in the range of 80 to 90 per cent that their Government had authorities responsible for the following functions: data-related privacy and confidentiality, technological and statistical standards, governance policies and management practices, and partnerships and collaboration with national statistical systems and civil society. However, there was less coverage for the functions of data-related work across Government, security of data systems and data exchange, data and statistics readily findable and accessible, user engagement, quality and interoperability, and good practices for use and reuse, in which national statistical offices reported between 70 and 75 per cent of their Governments having authorities responsible. Against this list, the majority of national statistical offices identified themselves as a responsible authority in Government for those roles, either solely responsible or jointly with another organization in Government.

8.   National statistical offices were asked whether their responsibilities had narrowed, remained unchanged or expanded for the same list of data stewardship functions. The majority of national statistical offices reported that their role had expanded for those functions, except over security of data systems and exchange and privacy and confidentiality, which had remained the same for the majority of respondents. Very few national statistical offices reported any narrowing of responsibilities with those functions. More than 60 per cent of national statistical offices reported that responsibilities had expanded greatest in data-related work across Government, technological and statistical standards, and partnerships and collaboration with national statistical systems and other stakeholders.

9.   Perhaps the most significant quantitative result across the questionnaire was that 80 per cent of national statistical offices that reported increased data-sharing and reuse as the success measure for data stewards. This was a prominent result because it was markedly higher than the next most popular response from the measures given

in the list, which were increased trust (65 per cent), active engagement between producers and users (65 per cent), more data-driven and statistics-driven policymaking (63 per cent), increased resources allocated (58 per cent), increased availability and timeliness (46 per cent), increased data literacy (39 per cent), low incidence or impact of data privacy breaches (33 per cent), increased citizen engagement in surveys and censuses (15 per cent), other measures (11 per cent) and increased use of statistics by media (9 per cent).

10.   The final quantitative question on links between the national statistical office and municipal/city showed that 77 per cent of national statistical offices had either an existing organizational function or formalized process or held ad hoc meetings with municipal or city government data stewards. While 9 per cent had rare or no engagement, the link was not relevant for 10 per cent, and the rest did not know. A background document accompanying the present report will contain further details on the questionnaire results.

III.   Updates from the workstreams

A.   Data governance workstream

11.   This workstream, which is co-led by Statistics Poland and the World Privacy Forum, provides guidance to national statistical offices regarding data governance and legal frameworks to assist them in transforming and potentially expanding their role as data stewards. In its initial phases of work, the workstream sought to probe the boundaries and interaction of existing legal frameworks and statistical governance practices through an examination of various country case studies, including an analysis of how major data protection frameworks such as the General Data Protection Regulation had been adopted around the world,[7] and to understand current and emerging practices in relation to official statistics. Canada, Mauritius and New Zealand comprised the first case studies, with an analysis of the legal and regulatory frameworks in place in those jurisdictions, as well as an analysis of those jurisdictions' forward-looking proposals and new uses of statistical data and other data.

12.   Initial qualitative results exhibited meaningful regional variations in approaches, suggesting that geographical, economic and other contexts could hold a significant role affecting the practical implementation of data governance practices for statistical uses. In the background report for 2022, the workstream concluded that "data governance has proven to be highly contextual in the modern world".[8] Follow-on work to clarify and deepen understanding of contextual differences in data governance and stewardship approaches among national statistical offices has been accomplished in three distinct phases. First, additional interviews and discussions were conducted to further understand and crystallize the initial indicators of the nature and type of impact that regional differences had on practical data governance in the modern statistical context. Second, additional legal framework analysis was conducted to further clarify the impact of national statistical organizations' exemptions from large regulatory frameworks such as the General Data Protection Regulation. This included interviews to clarify perceived negative impacts of

---

[7] Pursuant to General Data Protection Regulation, chapter 9, article 89 (see https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679). European member States may have additional derogations regarding statistical work. See General Data Protection Regulation derogations tracker, available at https://gdpr.lw.com/Home/Derogations.

[8] See Working Group on Data Stewardship, "Preliminary results from the Working Group on Data Stewardship" (March 2022).

exemptions. Lastly, the construction, fine-tuning and analysis of the global consultation questionnaire regarding the perceived role of national statistical offices as data stewards was an essential element of sharpening understanding of the issue.

13.   With the global consultation results now available, the workstream's focus has shifted to refining and substantiating new understanding emerging from the questionnaire responses. There are two areas of the questionnaire with specific importance for this workstream. First, the results indicated that increased data use and reuse was a top criterion of how a national statistical office defined the successful implementation of a data stewardship approach. This finding was clear and statistically significant and has provided a meaningful avenue to explore further. The second result indicates clear regional adaptations to data governance approaches, which supports and clarifies the workstream's early findings.

14.   The revised workplan for the workstream includes the conduct of additional qualitative interviews aimed at refining understanding in two key areas: (a) regional interpretations and implementations of data use governance methods, regulations, and tools; and (b) emerging regional best practices concerning data use and data-sharing across these same regions. To address the first question, the workstream will plan a series of qualitative interviews across a sampling of additional regions to further understand differences and to learn about national statistical offices' current and emerging best practices regarding data-sharing. To further address the question of data sharing beyond new interviews, the workstream is also researching a variety of examples of best practices regarding data-sharing by national statistical offices. In one example, some national statistical offices are utilizing GitHub repositories to share their data sets and create the possibility of broader dissemination. Belgium, Latvia, Mongolia, New Zealand, Poland, Thailand and the United Kingdom of Great Britian and Northern Ireland are among the countries whose national statistical offices have utilized GitHub to share data according to best practices and make the data readily available in an appropriate manner. The focus of this next tranche of work will be to compile a broad and diverse array of use cases and best practices to share with the Commission.

B.   Equity and inclusion workstream

15.   The equity and inclusion workstream is co-led by the Departamento Administrativo Nacional de Estadística and the Global Partnership for Sustainable Development Data to provide guidance to national statistical offices to ensure that an equitable and inclusive approach to data use is adopted in order to achieve the goal of the 2030 Agenda of leaving no one behind. In 2023, the workstream has updated the briefing document developed in 2022 with the findings from the global consultation questionnaire and from the two country case studies, received from Canada and Colombia. The updated briefing document will help to identify the key priorities and progress on this topic. It will also act as a guide to a series of consultations and sprints that the workstream will carry out with countries in 2024. This will help to further refine and share knowledge on equity and inclusion in implementing the 2030 Agenda and in guiding data stewardship across the entire data value chain. The sprints will further refine the work of the workstream beyond 2024.

C.   Sharing and collaboration workstream

16.   The sharing and collaboration workstream is co-led by the Departamento Administrativo Nacional de Estadística and the Instituto Nacional de Estadística y Censos (Argentina) with the aim of developing a better understanding of issues

relating to access, sharing and collaboration within the data ecosystem. The workstream is diligently working on a document aimed at fulfilling the initially established deliverables. These deliverables are intended to provide comprehensive guidelines that address the increasing need for sharing and collaboration, with an emphasis on the benefits of cooperation. The deliverables also make an attempt to give visibility to the importance of sharing and collaboration when creating value in data, especially through the promotion of transparency, creating synergies among organizations and promoting collaboration in innovation and discovery.

17.   Furthermore, the document is intended to elucidate the current challenges pertaining to data-sharing and collaboration. It encompasses issues such as privacy concerns, varying regulations across regions, the legal responsibilities of national statistical offices, technical challenges in achieving interoperability and the assessment of diverse international scenarios. The deliverables also endeavour to establish protocols and standards for data-sharing, offering real-world examples of collaboration agreements. Moreover, the document is intended to share insights gained from the negotiation processes for data-sharing and identify the instruments best suited for data-sharing needs. Building on the resources accumulated by the workstream over time, the document is intended to provide descriptions of real cases, an analysis of the factors facilitating successful data exchange, and the lessons learned from these experiences. Lastly, the deliverables from the workstream will include a guide and templates for negotiating data-sharing agreements with members of the data ecosystem. In addition, this group extends an invitation to countries to participate in the workstream, with the objective of ensuring that all corners of the globe are represented. The aim is to move closer to a comprehensive global embodiment of the data stewardship concept, with a commitment to inclusivity and an acknowledgment of the varying starting points that each country may have in its approach to data stewardship.

D.   Data stewardship and the city data agenda workstream

18.   This workstream was led by the Sustainable Development Solutions Network's Thematic Research Network on Data and Statistics. The workstream linked the Working Group and those engaged in the smart cities initiative to foster knowledge exchange, identify areas of mutual interest and suggest remedies in these areas to improve and connect data stewardship at the national and city levels. The workstream concluded its work during 2023; the other workstreams will continue to make connections and use references to the earlier work of this workstream, where necessary.

E.   Overall conceptual framework on data stewardship workstream

19.   This workstream is co-led by Statistics Poland and Open Data Watch to establish a common understanding of what the concept of data stewardship covers, recognizing that there are various interpretations of data stewardship and differences in the terminology used across data communities. The goal is to provide guidance on the role of national statistical offices as data stewards supporting and supplementing their basic function of producer of official statistics. This workstream strives to reconcile regional definitions of data stewardship to develop a global conceptual framework that describes data stewardship on the basis of a common terminology. This will ensure coherence across the other three workstreams in the Working Group and across the various regional data stewardship activities.

20.   The workstream collaborated with other Working Group members to revise, finalize and deploy a global questionnaire to gather comprehensive feedback on the roles, responsibilities and indicators of success of a data steward. The analysis of the globally and regionally representative responses has served as a critical input for a preliminary checklist of functions of a data steward (see annex II to the present report). Most respondents emphasized that a key role of a data steward is to facilitate and encourage data-sharing and reuse. This can be achieved by building trust in data and promoting data-sharing and use in decision-making, safeguarding data quality, facilitating greater collaboration, coordination and data integration across the national statistical system, maintaining citizen confidentiality and data security, and implementing strong data management and capacity-building practices. However, the workstream reiterates that data stewardship is a context-sensitive concept and that any checklist or framework must be adaptable to ensure that national statistical offices can use it to best serve their country's specific needs.

21.   Following a more detailed analysis of the global consultation results, this workstream will develop a more complete checklist of functions of the data steward, which will inform the creation of an implementation framework and further guidance. The results will also be used to create a global glossary of terms and reconcile the outputs of this Working Group with any relevant outputs from the five United Nations regional commissions, including the glossary from the Conference of European Statisticians/Economic Commission for Europe task team. These deliverables will be disseminated broadly through webinars prior to and during the fifty-fifth session of the Commission. These webinars will serve as an opportunity to raise awareness and increase understanding of the checklist and gain feedback and input from relevant stakeholders. This may require additional consultations with key stakeholders and the development of a compendium of case studies of how data stewardship is operationalized within various data and government systems.

IV.   Action to be taken by the Statistical Commission

22.   **The Commission is invited to take note of the present report.**

Annex I

Workplan (January–December 2024)

I. Governance and legal frameworks workstream

1. Promote the role of official statistics in the national and supranational data governance systems, which can effectively contribute to the development of efficient and effective data ecosystems, in particular in the coordination of national information systems.

2. Encourage the possibility of developing the ethical guidelines for data stewardship, in particular in obtaining data from new data sources (including private and big data) and new models of dissemination vis-à-vis the data sources (origination), and of incorporating the idea of artificial intelligence and machine learning guidance into data governance, steward processes and roles, to foster trust.

3. Support the development of the capacity of national statistical offices to fulfil the role of data stewards, taking advantage of statistical infrastructure (nomenclature, standards, classifications, information models and statistical frameworks) and recent achievements and increased experience in using new data sources and ecosystems.

4. Promote the development of cooperation with other relevant stakeholders of national and supranational data ecosystems by creating an observatory on data stewardship, inclusive of best practices in the realm of data governance and examples of data stewardship in national statistical offices' activities.

5. Ensure that best practices are used concurrently with the framework of the Sustainable Development Goals to nurture progress.

Deliverables

- Development of three areas: ethical guidelines for data stewardship; an observatory on data stewardship; and development of of national statistical offices' capacity.

- Promotion of and support for the development and implementation of national capacity-building strategies through regional workshops in the Africa and the Asia and the Pacific regions, in collaboration with the United Nations regional commissions and other groups, such as the Expert Group on National Quality Assurance Frameworks.

II. Equity and inclusion workstream

6. Identify specific research questions to shed light on thematic and sectorial dimensions of the equity and inclusion agenda (Indigenous communities, gender and intersectional approach), to support various national statistical offices in understanding how they can build their own approach as data stewards while achieving relevance by meeting users' expectations, as a cornerstone of statistical quality assurance. These research questions will guide the case studies.

7. Systematize learning from various national approaches that have been implemented to foster equity and inclusion, strengthen national statistical offices' role as stewards, and foster the empowerment of and collaboration with vulnerable communities through improved governance schemes.

8. Develop guidance on how national statistical offices can promote equity and inclusion along the data value chain to strengthen their role as stewards, recognizing

the realities under which this promotion must be undertaken, and the specific expectations that users have regarding the national statistical offices' role and their own.

Deliverable

- Fully defined research questions on equity and inclusion (thematic and sectorial dimensions), learning from national approaches and guidance on promoting equity and inclusion along the data value chain.

III. Sharing and collaboration workstream

9. Diagnose the various international scenarios in which sharing and collaboration are being discussed under the data stewardship concept.

10. Document the best practices and lessons learned by national statistical offices and other stakeholders while negotiating and achieving data-sharing agreements with public and private actors.

11. Provide guidance on the instruments that better suit data-sharing and collaboration needs and identify research questions regarding the enabling factors for collaborative environments with multiple stakeholders under the data stewardship concept.

Deliverables

- Development of a set of guidelines for negotiating and achieving data-sharing agreements with public and private actors.

- Exploration of the types of instruments that best suit the national statistical offices' data-sharing needs and through an inventory of resources, guides and templates.

IV. Overall conceptual framework on data stewardship

12. Develop a robust and adaptable framework on data stewardship.

13. Issue guidance on the role of national statistical offices as data stewards within Government and data systems.

Deliverables

- A checklist to identify the needs and responsibilities of data stewards.

- Creation of a framework for data stewardship within the national statistical system, including formulation of commonly agreed glossary of terms (with the Conference of European Statisticians/Economic Commission for Europe task team).

- Promote and develop the overall conceptual framework on data stewardship through a webinar series.

V. Meetings

Ninth meeting: first quarter of 2024

Tenth meeting: second quarter of 2024

Eleventh meeting: third quarter of 2024

Twelfth meeting: fourth quarter of 2024

Annex II

Preliminary checklist of functions of a data steward

Data stewards encourage data-sharing, use and reuse across the national statistical system, including administrative data, by performing the five related/ interlinked functions outlined below.

Building trust in data and promoting data-sharing and use in decision-making	*Safeguarding data quality*	*Facilitating greater collaboration, coordination and data integration across the national statistical system*	*Maintaining citizen confidentiality and data security*	*Implementing strong data management and capacity-building practices*
• Engaging data producers, users and other relevant stakeholders in data-related activities	• Maintaining the accessibility, timeliness and coherence of data flows	• Coordinating the production and integration of official statistics within the national statistical system	• Applying guidelines and practices for the privacy and confidentiality of data across the national statistical system	• Developing governance policies and management practices for the national statistical system and broader data ecosystem
• Building the data literacy skills of data users	• Ensuring that data adhere to relevant international and national quality standards	• Promoting data-related activities across government agencies	• Ensuring the security of data systems and data exchange processes within the national statistical system	• Overcoming capacity limitations of the various parts of the national statistical system to integrate data produced by non-State entities
• Fostering partnerships and collaboration opportunities between the national statistical system and civil society organizations	• Establishing and maintaining data interoperability and facilitating data integration	• Addressing perceived boundaries between non-State-owned data and official statistics		• Establishing and adhering to technological and statistical standards
	• Implementing principles of open data by default so that data and statistics are readily findable and accessible			

———————————