



Economic and Social Council

Distr.: General
2 December 2021

Original: English

Statistical Commission

Fifty-third session

1–4 March 2022

Item 3 (b) of the provisional agenda*

Items for discussion and decision: data stewardship

Report of the Working Group on Data Stewardship

Note by the Secretary-General

In accordance with Economic and Social Council decision 2021/224 and past practices, the Secretary-General has the honour to transmit the report of the Working Group on Data Stewardship. The report provides a summary of the recent activities of the Working Group related to its five workstreams and details the working definition for data stewardship established by the Working Group.

The Commission will also have before it a background document on the Working Group on Data Stewardship.

Action to be taken by the Commission is set out in paragraph 33 of the report.

* E/CN.3/2022/1.



Report of the Working Group on Data Stewardship

I. Introduction

1. The Working Group on Data Stewardship was established by the Statistical Commission at its fifty-second session. 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. In its decision 52/103, the Commission tasked the Working Group with exploring ways forward to develop guidance for national statistical offices on approaches to data stewardship.

2. There are currently 15 national statistical offices involved in the workstreams of the Working Group, from Albania, Argentina, Australia, Canada, Colombia, Ecuador, Germany, Latvia, Lithuania, Malawi, the Netherlands, New Zealand, Norway, Poland and Sweden, and 5 organizations from other data communities, namely, the Global Partnership for Sustainable Development Data, Open Data Watch, the Sustainable Development Solutions Network, the Economic Commission for Europe and the World Privacy Forum. The current Co-Chairs are Juan Daniel Oviedo Arango, Director General of the National Administrative Department of Statistics of Colombia (DANE-Colombia), and Dominik Rozkrut, President of Statistics Poland.

3. In addition, 15 national statistical offices have shown an interest in this work and remain on the invitation list to attend overall Working Group meetings, from Austria, China, Estonia, Finland, Indonesia, Kuwait, Malaysia, Mongolia, the Philippines, the Republic of Korea, Rwanda, Spain, the United Kingdom of Great Britain and Northern Ireland, the United Republic of Tanzania, Viet Nam and Zimbabwe.

II. Defining the workstreams

4. During the first meeting of the Working Group, on 29 April 2021, members expressed their views regarding the different dimensions of data stewardship and agreed to define five workstreams: data governance; equity and inclusion; sharing and collaboration; data stewardship and the city agenda; and the overall conceptual framework on data stewardship. During 2021, the overall Working Group met twice, and the individual workstream groups met two to three times each.

5. In establishing the working arrangements for the Working Group, the Co-Chairs proposed some changes to the terms of reference involving the establishment of an annual workplan to be presented to the Statistical Commission each year for approval. The amended terms of reference are contained in annex I.

A. Data governance workstream

6. This aim of this workstream, which is co-led by Statistics Poland and the World Privacy Forum, is to provide guidance to national statistical offices regarding data governance and legal frameworks to assist them in transforming and potentially expanding their role as data stewards. The guidance will share the experiences of national statistical offices in dealing with governance issues and developing legal frameworks to determine where new boundaries exist between data and statistics, where the key risks are, what a broadened horizon might look like and what is most useful in different contexts and at different starting points. The workstream will also

foster a broader conversation about pressing issues relating to governance and legal frameworks for data and statistics.

7. The workstream will address gaps across the Working Group by gaining a fuller understanding of the governance and relevant legal ecosystems of countries and regions through shared information, case studies and, in some cases, collected data. While much work has already been done on governance and legal frameworks, the past 5 to 10 years have seen a considerable expansion of new legal frameworks, as well as significant changes in data governance practices and architectures. These changes have created new areas of opportunity and tension for national statistical offices. The workstream will directly address these topics, specifically for the national statistical office context.

8. The workstream includes gathering data from members of the Working Group on existing and forthcoming documents related to data governance, at the national statistical office, country and supranational (e.g., European Union) levels. There is movement among some countries and national statistical offices toward crafting general guidance on data ecosystem governance, in some instances encompassing guidance on artificial intelligence and machine learning, which the workstream group is actively compiling. The group is working toward the creation of a glossary of data governance terms used by national statistical offices to enable clearer communication among such offices and to facilitate an understanding of different data governance models and how they are used by government institutions with different data governance responsibilities. Finally, the workstream group is collecting key data governance use cases from national statistical offices and analysing them to ascertain and catalogue best practices.

B. Equity and inclusion workstream

9. The aim of this workstream, which is co-led by the National Administrative Department of Statistics of Colombia and the Global Partnership for Sustainable Development Data, is to ensure that an equitable and inclusive approach to data use is adopted in order to achieve the pledge made in the 2030 Agenda for Sustainable Development of leaving no one behind. As such, this workstream seeks to expand the outcome of better access set out in the background paper presented to the Statistical Commission at its fifty-second session¹ by examining the policies and practices of members of the data ecosystem to ensure inclusive and equitable data use and the role of government data stewards in fostering the inclusion of different communities throughout the data value chain.

10. The main deliverable of this workstream will be a briefing document, created through the compilation of case studies of different stakeholders, which highlights the main existing trends, but also raises questions for further research on the matter. The workstream group will jointly define the scope of equity and inclusion within the framework of data stewardship through the construction of a common understanding of the challenges national statistical offices face in fulfilling their mandates; invite members to participate in the debates and conceptualization processes needed to build the briefing document and collect countries' experiences in overcoming these challenges for inclusion in the document.

11. The first action taken by the workstream group to promote the participation of other members was to draft a short document defining the relationship between data stewardship and equity and inclusion. In the document, it is stated that countries'

¹ Available at <https://unstats.un.org/unsd/statcom/52nd-session/documents/BG-3a-DataStewardship-E.pdf>.

ability to overcome the barriers that prevent vulnerable groups from benefiting from data is deeply related to the principle of leaving no one behind. In this context, the different members of the national statistical systems, and particularly the national statistical offices, could work on the operationalization of data equity and inclusion within their governance and planning arrangements.

12. Although there is not a single answer as to how to operationalize this mandate, there is a role for national statistical offices to act as data stewards, in many cases alongside others who steward data in the national data ecosystem. Members of the data ecosystem are evolving to become not only users of data, but also producers of data, sometimes playing roles that were only possible for national statistical offices in the past. Given this situation, the national statistical offices could expand their role to include overseeing and ensuring that the data provided by other stakeholders complies with pre-established standards of equity, inclusion, quality, relevance, impartiality, misuse prevention and confidentiality, thereby broadening their functions to include stewardship.

13. Taking this into account, the members of the workstream group defined three main challenges to fully guaranteeing equity and inclusion in the statistical processes within the data stewardship role of national statistical offices. First, the national statistical offices need to identify barriers and enablers to encourage widespread data use by different stakeholders in society. They also need to focus on issues such as data literacy and the development of an ethical framework that allows all types of users to navigate an increasingly decentralized data ecosystem. Second, it is paramount that increased data use results in inclusive and equitable policies and practices. National statistical offices, in coordination with decision-makers and public institutions, should address the limitations of different types of data to inform public policy. Finally, national statistical offices need to define their approach to fostering the inclusion of different communities along the data value chain and in data governance, generating inclusive data for both policymakers and the communities themselves. To attain a world without poverty and discrimination, the development path must be built and travelled together, particularly with the most vulnerable. There should be broad participation in the collection, processing, dissemination and governance of data to ensure equity and inclusion in statistical processes and improve outcomes for all.

C. Sharing and collaboration workstream

14. The aim of this workstream, which is co-led by the National Administrative Department of Statistics of Colombia and the National Institute of Statistics and Census of Argentina (INDEC-Argentina), is to contribute to the understanding of sharing and collaboration within the data ecosystem and to ascertain how national statistical offices can coordinate their activities with other stakeholders of the data ecosystem to improve the use of data in society. While acknowledging the existence of different legal frameworks across the globe, the workstream group did not consider the barriers this could raise, focusing instead on defining the data ecosystem, identifying the participating stakeholders and proposing sharing and collaboration mechanisms for better understanding and demystification of the concept of data stewardship to allow the expansion of the concept to all regions.

15. The workstream delivered the following main outputs: a document containing an analysis of different scenarios for sharing and collaboration between national statistical offices and other stakeholders within the scope of the data stewardship concept and a document that synthesizes the different experiences of the member countries of the workstream group with respect to public and private sector

agreements and data collaboration, providing insight concerning the general challenges that are being faced in order to focus further work.

16. The documents focus on addressing the following challenges and questions concerning how to improve collaboration:

(a) Are the data coordination bodies part of, or separate from, the national statistical system and national statistical office coordination bodies? How are the coordinating bodies related to the national statistical system?

(b) What role, if any, should the national statistical office play in professional networks and secondments in the national data ecosystem?

(c) What challenges are faced in the negotiation of data-sharing agreements with different partners in a data ecosystem, and can lessons from national statistical offices be applied in the wider data ecosystem?

(d) What should the format be of a standard data-sharing agreement to be applied to the new data ecosystem?

17. Members of the workstream group discussed the point of departure for the general concept of data stewardship. Members raised the issue of the different starting points of national statistical offices, as defined by their institutional and regulatory environments, as well as by their missions and roles in the data ecosystems of their respective countries. In addition, members highlighted the fact that while some regions and countries have been advancing fast in adopting a data stewardship approach, others have not agreed on its definition. It is therefore unclear whether an environment for data collaboration has been established.

18. One essential issue on which all the members expressed agreement was that cooperation is a must to maximize the use of alternative sources of data, including administrative records from other official entities, as well as sources from new producers in the private sector, while one of the main challenges is the establishment of a common language to foster data-sharing and collaboration. This issue could be addressed by using international standards, defining legal frameworks and expanding data literacy programmes.

19. Members were asked to provide case studies that provided guidance on the tool kits used to negotiate data exchange with other stakeholders in the national data ecosystem and how confidentiality is maintained within the agreements. All of the case studies were included in the inventory of resources, which was one of the deliverables of the workstream. The workstream group will strive to strengthen the inventory of resources with the experience of other members, refine the template for data-sharing agreements and address the questions highlighted in the background paper submitted to the Commission at its fifty-second session¹ to improve sharing and collaboration within the national statistical system.

20. Finally, members were asked to provide answers regarding the main challenges encountered and best practices identified in negotiating and achieving public and private sector agreements and data collaboration. The outcome of this discussion will be organized into a document to show the different experiences within the workstream membership of public and private sector agreements and data collaboration.

D. Data stewardship and the city data agenda workstream

21. This workstream, which is led by the Thematic Research Network on Data and Statistics of the Sustainable Development Solutions Network, links the Working Group and those engaged in the smart cities initiative, with the aim of fostering knowledge exchange, finding areas of mutual interest and suggesting solutions to

improve and connect data stewardship at the national and city levels. The workstream group developed a short briefing document with a case study and the results of one-on-one consultations and focus groups of high-performing and lower-performing countries.

22. The initial step taken by the workstream group was to develop a set of research questions that would drive the work programme, consistent with its role under the Working Group, as follows:

(a) Can a national-level data stewardship role be improved by drawing from experiences at the city level?

(b) What are the job descriptions, skills and capabilities of chief data officers at the city level and can national-level data stewardship draw from these city-level profiles?

(c) What are the options for establishing a practical permanent link between the work of the Statistical Commission and the smart cities initiative that would be mutually beneficial?

(d) Can the community of practice among smart cities be replicated at the national level to foster support and provide technical assistance consistent with the specific needs of national data stewards?

(e) Can the action-oriented efforts of city-level data stewards inform policy-driven approaches at the national level?

23. The process for addressing the research questions was as follows: a background literature review was conducted to identify the policy drivers for the establishment of chief data officers at the city level; interviews were conducted with city-level chief data officers to document their experiences and challenges; a case study was written on León, Mexico; and a conceptual framework was developed to organize the lessons from the interviews and create an assessment that easily translates to the national-level data steward role. The framework's components include an enabling environment, which includes support from the political and executive branches of municipal government and support from higher levels of government, along with the legal, regulatory and organizational structures under which city-level chief data officers operate; the capabilities of chief data officers, which comprise technical skills, the ability to work with and influence across the government and the ability to conduct outreach and external consultations; and the nurturing modalities that are the mechanisms that foster mutual learning, peer support and knowledge exchange.

24. Although this work remains in its early stages, some common observations are emerging. The city-level chief data officers spoke of the difficulty of overcoming data silos. There can be entrenched control of local data products or a lack of understanding about the potential benefits of making the data more accessible; the initial role of the chief data officer can accelerate or inhibit transformation. Some offices began in data registry management and compliance, which can create a rebranding challenge as they move towards a service-based role. Others began as incubators of innovative data analytics or in research supporting planning, which can make it easier for chief data officers to play a service provision role. It is important to use external expertise to support this agenda, as the offices of chief data officers often require help with issues related to cloud computing, data science analytics and data management, among other things.

25. In terms of fostering exchange of knowledge between national statistical offices and city-level chief data officers, interviews suggest that there is a need for an entity to actively pull these two communities together, with a focus on the development of capabilities and sharing of experiences.

E. Overall conceptual framework on data stewardship workstream

26. The aim of this workstream, which is co-led by Statistics Poland and Open Data Watch, is to establish a common understanding of what the concept of data stewardship covers, considering that there are different interpretations across different data communities and different uses of the terminology. The workstream strives to provide clarity on the conceptual framework and the boundaries of a common terminology for data stewardship. It will ensure coherence across the other four workstreams.

27. The workstream will define the boundaries for data stewardship for official statistics and provide examples that fall outside of those boundaries. A stocktaking of existing definitions was undertaken to gain a better understanding of the current landscape and identify commonalities within existing definitions. This exercise included a review of 34 documents from the private sector, multilateral organizations, government entities, academia and civil society organizations. It found that data stewardship was not clearly defined and was often discussed in vague and abstract terms. The interpretation of data stewardship ranges from complex definitions that include providing data, curating data, encouraging data use, facilitating data dissemination and creating regulations, to others that only mention managing data, with little description.

28. There are also variations in the term used for data stewardship in different languages, making it difficult for the term to be used and understood in a global context without a strong common definition. In French, for example, the term used is “intendance des données”, which corresponds directly to the English term. In Spanish, the term used is “administración de datos”, which highlights the role of data management but not the other functions of data stewards. Research on this subject, along with important inputs from the members of the workstream group, informed a piece on defining data stewardship written by the workstream Co-Chairs for the United Nations World Data Forum blog series.²

29. Based on the outcome of the stocktaking exercise and follow-up discussions, members emphasized that there was no one-size-fits-all approach to data stewardship. Members agreed that in both its definition and application, data stewardship must be context-specific and will depend on the possible roles of national statistical offices. With that in mind, any conceptual framework developed by the workstream group must be adaptable to countries with varying levels of capacity.

30. Some common elements were identified that must be included in the data stewardship framework. These include a common understanding of the data assets that are being considered. The assets can be categorized as people (data collectors, analysers and users and managers of statistical systems), technology (technical infrastructure) and processes (governance, laws, policies and procedures) within a country’s data ecosystem.

31. Effective data stewardship ensures that these elements work harmoniously to increase the trust in and the value, use and impact of data for the public good. Within this framework, the data steward provides oversight and guidance, reduces risk and increases collaboration across the system. However, where this role is situated within the data system and how its functions are carried out vary by country. The workstream will strive to build on this framework by collaborating with efforts by the Economic Commission for Europe to develop a common definition. This may require follow-up on what is being undertaken within various statistical systems, as well additional consultations with key stakeholders and the development of a compendium of case

² Available at <https://unstats.un.org/unsd/undataforum/blog/defining-data-stewardship/>.

studies of how data stewardship is operationalized within varying data and government systems.

32. The increased data demands arising from the adoption of the 2030 Agenda for Sustainable Development have led to a need for innovation and a widening data ecosystem with a rapidly growing range of data actors. As a consequence, there has been a need for the role of national statistical offices to evolve from data producers into coordinators and, finally, into data stewards. While the workstream group is not proposing a unique definition of data stewardship, it agreed on the need to develop a common framework that accounts for the cross-cutting nature of data, so that the important role of data stewardship in ensuring that data add value and are used for good can be better communicated.

III. Action to be taken by the Statistical Commission

33. **The Commission is invited:**

- (a) To take note of the present report on the work of the Working Group on Data Stewardship;**
- (b) To express its views on the scope of data stewardship;**
- (c) To express its views on the progress of the five workstreams defined under the Working Group;**
- (d) To approve the amended terms of reference of the Working Group contained in annex I to the present report;**
- (e) To review and approve the proposed workplan of the Working Group for 2022 contained in annex II to the present report.**

Annex I

Terms of reference of the Working Group on Data Stewardship

(Amended version of 21 October 2021)

I. Objectives

1. The primary objectives of the Working Group on Data Stewardship are as follows:

(a) To enhance knowledge and provide recommendations with respect to data stewardship that include approaches to data governance, equity and inclusion, sharing and collaboration and the overall conceptual framework for data stewardship and can be applied to a wider data ecosystem, taking into consideration the different starting points of national statistical offices;

(b) To promote the commitment of the statistical community and other stakeholders to working on the implementation of data governance in the national statistical system, in particular through the coordination role of national statistical offices;

(c) To promote sharing and collaboration with respect to data, whether public-public or public-private, allowing the regular use of public interest data for the generation of official statistics, while preserving confidentiality and privacy;

(d) To improve the performance of national data ecosystems through the expansion of methods and access inspired by a data stewardship approach.

II. Membership

2. The Working Group will consist of representatives from national statistical offices and other stakeholders, including international and regional organizations and civil society organizations. The Co-Chairs will be representatives of national statistical offices. The Working Group will comprise country representatives from all regions.

3. The Statistics Division of the Department of Economic and Social Affairs of the United Nations Secretariat will serve as the secretariat of the Working Group.

III. Organization and conduct of the work

4. The Working Group will work through electronic exchanges and periodic meetings, whenever possible. Where feasible, the meetings will be conducted in conjunction with other meetings being held throughout the year.

5. The Working Group will conduct its work in an open, inclusive and transparent manner, inviting experts, as appropriate, from civil society, the private sector and academia and other professional bodies to contribute their knowledge and experiences to the work.

6. The Working Group will continually review the different workstreams defined under the Work Group and ensure appropriate integration to advance with greater speed in the development of specific topics, and will convene forums and workshops in which specialized or emerging new topics are discussed in depth.

IV. Expected duration

7. The Working Group will operate under the auspices of the Statistical Commission and will report to it annually, seeking its guidance.
8. The mandate and functioning of the Working Group will be reviewed by the Statistical Commission as needed.

V. Proposed activities

9. In line with its mandate, the Working Group will carry out the following activities:
 - (a) Coordinate further consultations on the role of national statistical offices as data stewards, supported by the secretariat;
 - (b) Review the definitions, terminology and methodological developments and issues related to data stewardship, including data governance, equity and inclusion, sharing and collaboration and the overall conceptual framework, as well as how to conduct and institutionalize this approach through agreements, such as agreements for data-sharing, with an emphasis on safeguarding privacy;
 - (c) Develop recommendations on approaches to data stewardship for national statistical offices, including strengthening their capacities as data service providers.
10. Based on further guidance given by the Statistical Commission, the Working Group will update its activities and deliverables according to an annual workplan to be presented each year to the Commission for approval.

Annex II

Workplan (January–December 2022)

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, particularly in the coordination of national information systems.
2. Encourage the possibility of developing ethical guidelines for data stewardship, particularly with respect to 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 in the data governance and data stewardship 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, including nomenclatures, 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 activities of national statistical offices.
5. Ensure that best practices are used concurrently with the framework of the Sustainable Development Goals to nurture progress.

Deliverable: summary document articulating the process and recommendations

II. Equity and inclusion workstream

6. Identify specific research questions to shed light on thematic and sectorial dimensions of the equity and inclusion agenda (related, for example, to indigenous communities, gender and an intersectional approach) to support different case studies, allowing national statistical offices to understand how they can build their own approaches as data stewards, while achieving relevance by meeting users' expectations as cornerstones of statistical quality assurance.
7. Systematize learning from different national approaches that have been implemented to foster equity and inclusion, strengthen the role of national statistical offices as stewards and foster the empowerment of and collaboration with vulnerable communities through better 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 users have regarding the role of national statistical offices and their own role.

Deliverable: revised version of the briefing document

III. Sharing and collaboration workstream

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

10. Document the best practices of 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 types of instruments that best suit the 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:

- Inventory of resources, with the best practices currently carried out by members on the matter
- Guide and template(s) for the negotiation of data-sharing agreements with members of the data ecosystem, taking into consideration the nature of the stakeholders and the type of data shared
- Final document, including recommendations, research questions for the work ahead and refined outcomes of the other deliverables (inventory of resources and data-sharing agreement templates)

IV. Data stewardship and the city data agenda workstream

12. Expand interviews and case studies.

13. Revise the conceptual framework as needed.

14. Make recommendations drawn from the experience of cities to support effective data stewardship at the national level.

Deliverables: expanded briefing document with selected case studies, one-on-one consultations and focus groups. The studies will cover a mix of high-performing and lower-performing countries.

V. Overall conceptual framework on data stewardship workstream

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

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

Deliverables:

- Framework of data stewardship. This may include a checklist to identify the needs and responsibilities that data stewards may fulfil within national statistical systems
- Compendium of case studies which demonstrate the different roles of the data steward
- Focus groups and interviews with key stakeholders to socialize the proposed data stewardship framework

VI. Meetings

Third meeting: February 2022, in conjunction with the Statistical Commission (hybrid)

Fourth meeting: May 2022 (virtual)

Fifth meeting: September 2022 (virtual)

Sixth meeting: December 2022 (virtual)
