



Economic and Social Council

Distr.: General
7 June 2022

English only

Economic Commission for Europe

Conference of European Statisticians

Seventieth plenary session

Geneva, 20-22 June 2022

Items 2 and 8 (d) of the provisional agenda

Thirtieth anniversary of the Fundamental Principles of Official Statistics

**Topics of common interest to statistical and geospatial communities –
joint session with the Regional Committee of United Nations Global Geospatial
Information Management for Europe (UN-GGIM: Europe):
Developments in data stewardship**

Back to the basics: The concept of data stewardship and its linkages with the Fundamental Principles of Official Statistics (FPOS)

Prepared by Colombia and Poland

Summary

The document links the new and widely used concept of ‘data stewardship’ to the Fundamental Principles of Official Statistics (FPOS). While the official statistics community is rethinking their role, FPOS remain the key principles upon which they operate. The document aims to build understanding how the complex role of national statistical offices (NSOs) as data stewards is aligned with the FPOS.

The document is relevant to two agenda items at the Conference of European Statisticians 2022 plenary session: item 2 “Thirtieth anniversary of the Fundamental Principles of Official Statistics” and item 8 (d) “Developments in data stewardship”.



I. Introduction

1. With their creation and endorsement by the General Assembly of the United Nations, the Fundamental Principles of Official Statistics (FPOS) have been the centerpiece that enshrines the activities of producers of Official Statistics across the globe. These principles have been implemented in various ways over their 30 years of existence and have managed to stay relevant in spite of significant changes associated with technological, and socio-economic changes, which have spurred the proliferation of new data users, producers and an ever-widening wave of data needs.
2. One of the main approaches taken by NSOs to navigate through this changing process has been to change the way in which they see themselves, and how they should interact with others, which has led to the emergence of the “Data Stewardship” concept. The enlargement of the data ecosystem has challenged the official statistics community to rethink their role as the sole providers of statistical information, and steer to an ecosystem in which they are the stewards of the official statistics, and as such, next to a considerable enlargement of their initial remit, they have to acquire reinvigorated responsibilities, allowing to mainstream into a statistical system the key principles upon which they operate. The emergence of this concept, rather than having a conflicting nature with the FPOS, seeks to build on this system of beliefs to define a framework for collaboration and action.
3. Nevertheless, as there is not a common understanding of the concept, and in light of recent developments such as new data landscapes, changing data ecosystems, discussion about data versus official statistics, data revolution and many more data-driven phenomena, the members of the statistical community are leaning towards an understanding that Data Stewardship does not have a one-size-fits-all approach; NSOs are yet to understand their complex role as stewards complying with the FPOS. This paper aims to contribute to the analysis of how the Data Stewardship approach is aligned with the FPOS, based on the developments and discussions of the United Nations Statistical Commission Data Stewardship Working Group about its meaning and implications.
4. This document is divided into two sections. Firstly, the document builds on the progress of the UNSC Data Stewardship in defining the concept, its implications and applications. The second section uses the conceptual framework set in the first part to address, principle by principle, how the Data Stewardship approach can be understood as an enabler to uphold the FPOS and mainstream them into an evolving data ecosystem.

II. Data stewardship concept and scope

5. In 2021, during the fifty second session of the UNSC, the UN Working Group on Data Stewardship was established, through the decision 52/103 of the Commission¹. The WG was tasked with exploring ways forward to develop guidance for national statistical offices on approaches to data stewardship. The members of the Working Group include representatives from NSOs and stakeholders from other data communities, including international, regional and civil society organizations. The Working Group has been meeting for over a year, during which its members have expressed their views on the different dimensions of data stewardship and defined five workstreams: data governance; equity and inclusion; sharing and collaboration; data stewardship and the city agenda; and the overall conceptual framework on data stewardship.
6. From the progress done in each of these workstreams, some highlights have begun to emerge. These elements describe the data steward role as deepening the functions of NSOs, particularly in relation to their treatment and integration with other data stakeholders. Through a stocktaking exercise the fifth workstream established that “(...) data stewardship

¹ The decision of the Commission was enabled by the work carried out within the High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development developed the background document “*approaches to data stewardship*” which presented a stock taking exercise to foster a wide and inclusive discussion regarding 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, (...)" (Working Group on Data Stewardship, 2021, p.7).

7. The work carried within this framework seeks to contribute to the strengthening of this concept from different perspectives. The first workstream "Data governance workstream" aims 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. Regarding new data governance trends, the workstream revised other sources and country cases, concluding that in a changing data governance landscape there are clear indicators showing that NSOs are adapting to new governance activities and expanding roles as a balancing process between the uses of data and data protection. In this context the workstream pointed the following topics for the future of NSOs regarding data governance and data stewardship:

- Use of private sector data for government statistical purposes.
- Whole-of-government legal frameworks, particularly those that align cooperation with data authorities and NSOs.
- New or amended country-level Artificial Intelligence (AI) regulation or policy as an opportunity.
- Development contexts for NSOs.
- Understanding and documenting the prerequisites for an NSO to take on additional roles as a data steward.
- Relaxing the tension between NSOs and data protection authorities regarding the use of data for statistical purposes.

8. The second workstream, "Equity and inclusion", aims to ensure that an equitable and inclusive approach to data use is adopted to achieve the pledge made in the 2030 Agenda for Sustainable Development of "leaving no one behind". Members state that NSOs, as head of the National Statistical System (NSS), must guarantee the implementation of core values and principles not only in their statistics but in data provided by other stakeholders. NSOs need to improve better access to data, and to do so NSOs have three tasks:

- Identify barriers and enablers to widespread use of data by different stakeholders in society.
- Assure that the increased use of data results in inclusive and equitable policies and practices.
- Define their approach to foster the inclusion of different communities along the data value chain, generating and mainstreaming disaggregated data for both policymakers and the communities themselves.

9. The Sharing and Collaboration workstream aims to contribute to the understanding of cooperation 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. Following this approach, the group took as starting point the recognition of different national contexts and frameworks, to contribute to the demystification of the concept, highlighting that there is indeed no one size fits all for the implementation of the stewardship approach.

10. The fourth workstream "Data stewardship and the city data agenda" aims to foster knowledge exchange between the Working Group and those engaged in the smart cities' initiative, finding areas of mutual interest and suggesting solutions to improve and connect data stewardship at the national and city levels. One of the main conclusions of this workstream indicates that data stewardship activities at the city and national government levels could be mutually reinforcing.

11. From the experience at the city level, the workstream made suggestions for NSOs including to emphasize the importance of creating value from data; expand to an enabler of data quality; use the opportunity to better understand whole-of-government use of data to

improve NSO products, services, and the user experience. When developing data steward networks for cross-sector learning, NSOs should develop common tools to address global priorities, as well as work to foster a government-wide enabling environment to leverage the data as a strategic asset. NSO leadership must be capable technically, as communicators, and to influence across organizational boundaries. The suggestions made the workstream realize the clear need for an entity to actively pull the NSO and city Chief Data Officers' communities together to focus on capacity development and experience-sharing.

12. The fifth workstream "Overall conceptual framework on data stewardship workstream" aims 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. To achieve this objective, the workstream conducted a stocktaking exercise reviewing definitions of Data Stewardship existing in the private sector, multilateral organizations, government entities, academia, and civil society organizations. From this exercise the workstream obtained the following key words on the concept: provide data, curate data, facilitate data dissemination, establish trust, encourage data use, create regulations.

13. Members of the workstream agreed to establish that data stewardship must be context specific in both its definition and application. In that way, the workstream's work focused on creating an adaptable conceptual framework. To identify the elements that this framework should include the workstream launched a survey among its members, and the responses pointed that:

- Data stewardship includes governing the data ecosystem to improve the use and reuse of data for the public good.
- Data stewardship is the ethical and responsible creation, collection, management, and use of data.
- Data stewards, including NSOs, must facilitate coordination and cooperation between data providers and users.
- Data stewards ensure data protection, standardization, and quality to build and maintain trust.

14. From the exercises performed by the workstream and some follow-up consultations the common elements present in the data stewardship concept were synthesized as follows:

"These [the common element] include a common understanding of the data assets that are being considered. These assets can be categorized as people (producers of statistics, data collectors, analyzers, and users), technology (technical infrastructure), and processes (governance, laws, policies, and procedures) within a country's data ecosystem. Effective data stewardship ensures that these elements work harmoniously to increase 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". (Working Group on Data Stewardship, 2022, p. 36).

15. These initial elements reflect and interconnect the concept with the FPOS in different ways. The following chapter argues that data stewardship is deeply interrelated with the FPOS, breaking down the analysis on each of the principles. This exercise is not meant to be exhaustive, as we consider the concept definition to be an ongoing task.

III. Data stewardship and the Fundamental Principles of Official Statistics

16. The FPOS have been shaping statistics since its first adoption in 1992 by the Conference of European Statisticians. The wide range and global significance of the principles led them to be adopted by the UNGA in 1994, becoming thereby a commitment of national governments to hold the highest standards of the official statistics. FPOS have proven their importance for official statistics as they constitute the basis for enabling trust in

official statistics. The FPOS form a solid basis for all ethical and quality-related conceptual documents throughout the world. The European Statistics Code of Practice (CoP), The OECD Recommendation for Good Statistical Practice or the Principles governing International Statistical Activities can be named here as some of many examples of the transposition of the UNFPOS and their adaptation into various quality frameworks. FPOS constitute also the main quality axis in the UN Handbook for Statistical Organizations and in the UN National Quality Assurance Framework Manual.

17. Using these principles in official statistics contributes to the production of high quality official statistical information, which is crucial in analysis and informed policy decision-making processes in support of sustainable development, peace and security, as well as for mutual knowledge and trade among the States and peoples of an increasingly connected world, demanding openness and transparency (Economic and Social Council, resolution 2013/21). In this sense, Data Stewardship is set to become a strategy to improve NSOs ability to comply with the FPOS, furthering public trust in official statistics. This section aims to analyze the relationship between each of the Fundamental Principles and the Data Stewardship approach as described in the previous section.

A. Relevance, impartiality, and equal access

18. “Official statistics provide an indispensable element in the information system of a democratic society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information.”

19. We must begin by recognizing that national statistical offices are not the only producers of official data. The role of NSOs as data stewards consists inter alia of structuring and coordinating statistical production. By recognizing that national statistical offices are not the only data producers of official statistics, data stewards are also interiorizing the need to ensure that their statistics and those produced by other public entities meet the existing standards of relevance, impartiality and equal access. This role/function is not limited to merely reviewing compliance with the Fundamental Principles, and thus, involves fostering a common understanding of the data assets within the NSS (people, technology, and processes) in order to improve the quality of the data.

20. In regards with the sharing and collaboration functions, the data steward is expected to standardize the conceptual and methodological frameworks of official statistics production and ensure that the data outputs are available on an impartial basis. This coordination role also translates into benefits for the statistical production of the NSO. By accessing to new data sources, such as administrative records provided by public and private institutions, NSOs are able to enrich the relevance and timeliness of their outputs, either by complementing them with other quality sources, or using other sources as contrasting techniques. This can improve the relevance if their statistical operations, using it as a contrasting technique, integrating it to complement traditional sources or generating new information. Furthermore, relevance can be guaranteed through the steward's coordination, by identifying clear achievable needs, avoiding the duplication of information generated by different stakeholders of the data ecosystem, and promoting evidence-based public policy.

21. Moreover, taking into account the fact that statistics are more and more based on combined sources of data, including the so called “new data sources”, which is much in line with the data stewardship concept, ensuring that the use of new data sources is consistent with Principle 1 is a preference given to a multi-mode approach while collecting the information. Validation of information and constant crosschecks are indispensable in the broadening scope of official statistics. The more actors operate within the emerging data ecosystems, the richer they become but, at the same time, more trust is needed. This trust is engendered on the solid basis stemming from FPOS - the shared standards which ensure more transparency and stronger resilience.

B. Professional standards and ethics

22. “To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.”

23. Counting with professional considerations and professional ethics, as an NSO, helps to retain trust in official statistics as the public and users can rely on the fact that the processes carried out by the NSO are being conducted according to an existing normative framework and that are not moved by “political” drivers. When broadening functions towards a data stewardship role “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” (Working Group on Data Stewardship, 2021, p.4). These standards, regarding professional aspects of the data steward, guarantee that the personnel working in all producing entities of official statistics, particularly the top management, have the accurate profile to do so. This prevents the NSS from malpractices and assures the quality of the data while keeping away biased information.

24. Regarding ethical aspects, the pre-established standards ensure that, not only the NSO’s work is undertaken accordingly to requirements of equity, inclusion, quality, relevance, impartiality, misuse prevention and confidentiality, but that these requirements are fulfilled in all data provided by other stakeholders. In the understanding that NSOs guide the adoption of ethical and professional standards among the statistical environment, an opportunity to target equity and inclusion-related goals within official statistics is presented. This opportunity could help to make visible populations that are often, for financial or methodological reasons, underrepresented in official statistics, guaranteeing an inclusive approach in statistics (e.g.; indigenous communities, LGBTIQ+, among others). Concerning the standards, not only visibility is important, but also the involvement of population groups along the data value chain. And in doing the above, the data steward highlights the eventual need to transform standards to better collect, produce and disseminate data.

25. Following on the aspect of opportunities, it may be important to notice that nowadays, in the technology-driven world, where the boundaries of official and non-official statistics are blurred in the public’s opinion, the proliferation of new sources of data (e.g. big data) calls for other actors to be covered by the same protective layer of the highest ethical standards ensured by the FPOS: mass media, social media, private stakeholders, NGOs, citizens - they all constitute the modern sources of information, very much in line with the data stewardship approach. The openness to other stakeholders, beyond official statisticians, inherent to the data stewardship, is a powerful asset. Having them on board as regards professional standards and ethics can only result in benefits for us all, producers and users of data.

C. Accountability and transparency

26. “To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.”

27. Being transparent and open to scrutiny is fundamental to foster trust in official statistics. Since the adoption of the FPOS, NSOs have been improving the transparency of their work, publishing their metadata and microdata, defining publication calendars, receiving audits and external evaluations. However, as NSOs strive towards expanding their remit and becoming data stewards their responsibility is not only to make themselves accountable, but also making others accountable. Nowadays, other public institutions are producing statistical information and using administrative records to deliver useful data for decision making. In this context, data stewards are needed to promote transparency and become scrutinizers of information produced by other government agencies. Metadata and microdata publication, integration methodologies of administrative records, anonymization processes and the protection of personal data, among others, must be ensured and the NSOs’

expertise in the matter is paramount for this new paradigm. It may be interesting to observe that in the data-driven world there are already shared between different professions dealing with data. For example, statisticians and responsible journalists, in their pursuit for the truth, have common goals and foundations (very close to those set out in the FPOS), yet they are often far as regards different languages they speak. Therefore, the nature of data stewardship, with its emphasis put on being transparent through the mere fact of being understood, underlines the importance of communication in establishing closer links between different professional groups dealing with data for the common goal of informing and thereby empowering through the use of unbiased, solid, verified information.

28. Moreover, accountability and transparency are interlinked with the existence of professional and ethical standards, as it becomes a way in which these standards are shared with the wider audience.

D. Prevention of misuse

29. “The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.”

30. Stewards must promote data literacy to support the correct interpretation of the information. It is the steward’s responsibility to educate specialized and non-specialized audiences to better understand the scope and limitations of official statistics. NSOs have the responsibility to unlock the public value of data, not only for government decision-makers, but also to civil society organizations, private firms, and the public, particularly vulnerable communities. Misusing statistics may lead to misguided decisions, jeopardizing people’s agency and their trust in official statistics. This responsibility is also transferred to new official statistics producers, and stewards should also encourage them to promote data literacy.

E. Sources of official statistics

31. “Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.”

32. The Fourth Industrial Revolution has provided us with new data processing technologies that have allowed us to expand the sources from which we can draw statistical information. Our role as coordinators of national statistical systems demands that we encourage collaboration schemes that stimulate access and integration of different data sources to improve the timeliness, quality, costs and the burden on respondents. In the role of data stewards that means a twofold commitment. On one hand, access to so called non-traditional data should be widened. Those sources are heterogenous, as they include administrative records (in many cases already accessible to NSOs by law), but also big data with all its variations (sensors data, satellite images, social media data, web scrapping data) to end with privately held data, owned by businesses, which represent a valuable source of information and efforts are being made worldwide to ensure a free of charge access to them for the official statistics. The overall goal of combining data sources, next to enriching the traditional ones, is to produce new, more robust and less expensive statistical information. This means that fostering the use of more statistical sources, NSOs should be playing an important role in shaping data governance to facilitate sharing and collaboration – in other words they should assume the responsibilities of data stewards, which is an important aspect of the commitment they make.

33. On the other hand, a focus should be made on empowering other producers into using new sources of information, with the view of achieving both: burden reduction and higher quality of the outputs. The use of new processing techniques to manage and curate administrative information or to integrate it with other sources for missional tasks can be useful. Capacity-building to benefit from new sources will help promoting the FPOS compliance by other producers.

F. Confidentiality

34. “Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.”

35. Confidentiality plays a key role into maintaining trust, which guarantees the permanence of the data sources and the users’ reliance in official statistics. In this sense, production practices that comply with confidentiality standards are vital, and it is the task of the data steward to ensure that these good practices, particularly anonymization techniques, are fully obeyed by all the entities that produce official statistics, and by all the sources that NSOs use.

36. In this scenario, data stewardship implies data protection to maintain trust, which in turn means using the data collected only for statistical purposes, preventing that the data and information of the informant are used in contexts other than those to which they consented.

G. Legislation

37. “The laws, regulations and measures under which the statistical systems operate are to be made public.”

38. Formal governance schemes contain the mandates of any public institution and, therefore, they shape data ecosystem stakeholders’ expectations. From the data stewardship perspective, the NSOs are not only the leaders of the national statistical systems, but also are also the ones who are called to promote statistical legislation dispositions.

39. Additionally, as it is impossible for the legislation to foresee all possible interactions, the data steward should be keen to be effective in the creation of governance and coordination mechanisms suited to their mandate and national contexts, allowing to exercise their role, while also ensuring a transparent and open approach to the discussions that are being held. The FPOS are legally mainstreamed into national contexts through national legislations.

H. National coordination

40. “Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.”

41. As seen throughout this paper the concept of data stewardship revolves around the coordination role of the statistical system, which, following this principle, is essential to achieve consistency and efficiency. Coordinating involves ensuring the harmonious work between the ecosystem’s elements, guiding and creating collaborations, and assuring the compliance with standards. To be able to execute this role a national data strategy and a clear legislation with the explicit role of the NSO as data steward would be considerable assets.

42. National coordination as part of data stewardship assures the efficiency of the data and its processes, guaranteeing its relevance, promoting its integration and avoiding its duplication. Thus, national coordination as an operative skill of the data steward implies linking the data produced by state entities in a twofold sense, firstly by creating relationships between official entities at the national level, and secondly by relating entities in the national level with those in subnational levels. To get to that point, data stewards gather efforts in stocktaking and communication, and are able to influence other data stakeholders in order to raise awareness about the importance of data as a strategic asset in the decision making.

I. Use of international standards

43. “The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.”

44. This principle advocates for the use of international concepts, classifications and methods into the official statistics, which has been largely done by the national statistical offices in the understanding that in doing so they guarantee the compliance of their statistics with standardization requirements, which, as said before, are essential in order to retain trust. As data stewards, NSOs need to integrate these international standards into the practices of all the producers of data within data ecosystems to promote the consistency and efficiency of statistical systems at all official levels. For instance, data stewards could assist other entities providing data to comply with quality of international standards for the use of administrative records, or anonymization processes.

J. International cooperation

45. “Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.”

46. Considering that international cooperation in statistics contributes to the improvement of systems of official statistics in all countries, assuming the responsibility of data stewards means to create synergies for other actors of the national statistical systems to participate into bilateral or multilateral scenarios. In this way, data stewardship deepens into the mechanisms of sharing and collaboration to strengthen the national data ecosystem and its relationship with international actors. In addition, stewards themselves are called to share and collaborate among them in order to understand the best practices on data stewardship operationalization and adapting them into national contexts.

47. One of the on-going international debates resonates around the question on how the position of an NSO as head of the NSS is subject to change and evolve in the international scenario.

IV. Moving forward

48. Taking advantage of the celebration of the 30th anniversary of the Fundamental Principles of Official Statistics and recognizing that the scenario in which NSOs operate has shifted since the creation and the adoption of these principles, it is important to have a deeper look on how those changes and new approaches for remaining relevant are related to the core system of beliefs of the official statistics community.

49. While all principles can be related with the emerging approach of data stewardship, there are some which are more relevant in respect of redefinition of data ecosystem and different data-related functions resulting from it. Having a common understanding on what are the principles is not minor, as they become the glue that allows the community of official statistics to be a dynamic, collaborative and innovative group, and allow us to jointly build our position to discuss with other communities. Proof of this element can be seen on the on-going and strengthened relation between the statistics and geospatial relation, which started as something casual and has become an “unescapable marriage”, and now the idea of having statistics without them being framed in a specific geographic setting has become obsolete. It is then key to understand the huge tool that we have, and to use it to establish solid partnerships with other stakeholders and communities.

50. Nonetheless, just recognizing the usefulness of the FPOS for the NSOs in their process of interiorizing and mainstreaming their role as data stewards is not sufficient, as such in order to uphold their value and relevance, it is important to analyze which investments should be made and prioritized for this to happen.

51. Both data stewardship and the FPOS are concepts stemming from the international sphere, this paper’s aim was to show interrelations between the two which persist to be strong, proving thereby the unchanged relevance of the FPOS. However, those two concepts are products of their respective times. And as much as FPOS have been created using the institutional approach, and they have set the basic rules under which the so-called “data revolution” is functioning, a question arises whether they need to adapt to the changing environment. In an era of changes like appearance of new data sources, more focus moved

to data rather than statistics, user-centric approaches, dramatic pace and avidity of information, changes of fundamental notions and concepts, irruption of fake news, post-truths et. al, - all these happenings make statisticians reiterate the question about pertinence of FPOS as a commitment of a group of people. The first query is “if this group is sufficient”? In presence of other data stakeholders, data stewardship concept appears as imminent. And as much as the NSOs should strive to build partnerships with non-official data producers and owners to make those data part of official statistical outputs, the corner values of the official statistics, such as quality, standards, professional independence should never be neglected, nor compromised.

52. The timelessness and relevance of the FPOS make them fit to the new data-driven reality. However, the proliferation of new data sources, new stakeholders and new concepts, such as data stewardship is a good opportunity to reinterpret the Principles, to set directions of their new possible explanations, which seem to be unavoidable in spite of their pertinence and universal character.
