Statistical Commission

Fifty-second session

1 – 3 and 5 March 2021

Item 3(a) of the provisional agenda

Items for discussion and decision: Data and indicators for the 2030 Agenda for Sustainable Development

Background document

Available in English only

Approaches to data stewardship

Prepared by the United Nations Statistics Division for the High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development (HLG-PCCB)

Approaches to data stewardship:

a background paper

(version as of 17 February 2021) [1]

Introduction

1. Following decision 51/120 at the 51st session of the UN Statistical Commission, the High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development (HLG-PCCB)[2] embarked on a stock taking exercise to prepare the background for a broader discussion on approaches to data stewardship by National Statistical Offices. As part of the work, the group considered a few case studies contributed by some members, other countries' National Statistical Office (NSOs) from outside the group, and other issue partners[3].

2. This background paper is prepared based on case studies of existing practices attributed to a data steward approach. The examples given in the different case studies are organised according to the scoping methodology (see Annex I) established at the beginning of this stock taking exercise. Using this method, data stewardship is associated with four outcome 'pillars'; governance, collaboration, methods and access which in turn are supported by underlying outputs, activities and inputs/ needs. The examples in the case studies were therefore organised under the relevant outcome and output in this scoping methodology.

3. Special thanks to the following organisations for the supply of these studies: the NSOs who provided the situation for their organisations in Argentina, China, Colombia, Germany, Finland, Ireland, Mexico, and New Zealand and the United Nations Economic Commission for Latin America and the Caribbean (UNECLAC) and SDSN TReNDS for the case studies on the position of a chief data officer separate from the NSO in Estonia, France, Mexico and in a sub-national context, the city of Los Angeles.

4. The work was supported by subgroup of the HLG-PCCB comprising of Argentina, Finland, Mexico, Saudi Arabia, State of Palestine, the United Kingdom of Great Britain and Northern Ireland and the United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP). The case studies considered in this first phase of the work and presented in this paper revealed a wide range of approaches to data stewardship with different levels of involvement from the NSO.

Background

5. As new data actors emerge in a changing data ecosystem and the demand for data and statistics ever increases, the role of National Statistical Offices is rapidly evolving. New technologies allow for an unprecedented level of data collection, integration and analysis on

[1] An earlier version of the paper dated 18 January erroneously stated an Open Data Institute was created in Germany, this is corrected to state it is being proposed.

[2] HLG-PCCB webpage: https://unstats.un.org/sdgs/hlg/.

[3] For more details of issue partners, see https://unstats.un.org/files/HLG_communication_concept_note-final.pdf.

human behaviour and overall societal trends. Data is now the basis of many business models. Those that build vast and ever-increasing databases of individual data have become the most valuable companies on the planet. They also profit, as well as nurture brand recognition, from using these data for statistical products and services. As these data are increasingly utilized and integrated into national data and statistical systems, there is a need to safeguard the use of these data and redefine the national data governance system. To discuss official statistics' position in this new data ecosystem, the topic of the High-level Forum on Official Statistics (HLF-OS) at the 51st session of the UN Statistical Commission (UNSC) asked the question "Data stewardship – a solution for official statistics' predicament?"

6. In a summary of the HLF-OS presented during the 51st session of the UNSC, the idea of data stewardship had different interpretations and connotations depending on the national context, expectations and institutional set-up. However, there was a consensus that the adoption of a data stewardship approach enables the statistical office to evolve from being a statistics producer to becoming a service provider that facilitates a joined-up approach to data and statistics across different data and statistics communities. As a result of the HLF-OS discussion and summary, the UNSC at its 51st session requested the HLG-PCCB via decision 51/120:

"to establish a **working mechanism** open to other stakeholders and **linked to existing initiatives** at global and regional levels to address the issue of **data stewardship and the role of statistical offices in the new data ecosystem**; the work should consist of a **stock-taking exercise** and **specific recommendations** on the **way forward**, to be provided to the fifty-second session of the Commission."

7. This paper responds to this decision and represents the requested stock-take through the collection of case studies on data stewardship categorised into four dimensions: better governance, better collaboration, better methods and capability; and better access and understanding. Finally, proposing the way forward as the creation of a new working group to further explore and develop the themes from this paper. The Terms of Reference of this group is contained in Annex II.

Better governance

8. The composition, structure and governance modalities of national statistical systems vary widely both among countries and to a lesser degree within countries. In many cases, this reflects the history of a specific country; how official statistics have become an important function of national administrations; how the **National Statistical Office (NSO)**, for which official statistics are its primary purpose has been established; and how the various organizations engaged in the official statistics of a country have evolved to respond to user needs, technological changes, as well as to institutional changes.

9. The historical international literature on the governance of official statistics, such as the United Nations series, *Handbook of Statistical Organization and Management*, has traditionally concentrated on standards for outputs, methods, definitions and terminology, and on the principles to govern all activities of official statistics rather than make recommendations on how to organise official statistics in each country. When the literature

has looked at such matters, it has focused on a debate around decentralised versus centralised systems for producing official statistics; and concluded slightly in favour of a centralised system.

10. In recent years, the literature expanded first to embrace the idea of the **National Statistical System (NSS)**[4] and still further, the idea of a **data ecosystem** where different actors interact with each other to exchange, produce and use data. The trend is to try and convey with these terms the complexity and decentralised nature of data production and use in the digital age and reflect in the era of social media platforms that a very different style of content communication has emerged; one which empowers the individual, demands a shorter attention span and where images and sound bites are often preferred.

11. This first section looks at the **data governance** arrangements many of the case studies demonstrated. One case study defined the term data governance as the strategy, principles and institutional policies for data management; supervision of data management processes and practices. The key elements in this section are therefore the leadership role of a chief data steward, the existence of a data strategy and the application of quality assessment and assurance.

Role of a chief data steward

12. While the chief statistician is usually appointed based on their professional competence to become the chief executive officer of the NSO and also often responsible for the professional integrity of the whole National Statistical System (NSS), the case studies showed the chief data steward role fell often into three types of arrangements:

    - Data stewardship is an expanded mandate and function for the chief statistician and therefore the NSO (Ireland, New Zealand)
    - The NSO contributes or partners in data stewardship activities but it is not their main mandate or function and no single entity holds this role (Argentina, Colombia, Finland and Mexico)
    - The chief data steward position and associated activities are carried out in another part of Government either politically appointed (France and Mexico between 2013 and 2018) or not (Estonia and the city of Los Angeles) and do not formally overlap with the NSOs mandate or functions

13. In the first arrangement, while the expanded mandate and function to be the data steward in Government were a positive development for the NSO, there was some trepidation in going too far beyond official statistics to fulfil an expanded data stewardship role. For example, one case study mentioned requests to assist in helping other bodies to set up business intelligence systems, data warehousing services, operational analytics and programme evaluations (econometric analysis). There was a general question how the NSO would respond as the demands for data stewardship evolved beyond statistical purposes and the risk of contamination to the NSO brand by association.

[4] the ensemble of statistical organizations and units (statistical agencies) within a country that develop, produce and disseminate official statistics on behalf of the national Government (and other levels of government). It is the responsibility of each country to define the scope of its NSS.

14. The case study from the NSO in Ireland mentioned some mitigating measures employed when working closely with policy making departments such as:

    - publishing the Memorandum of Understanding or similar document providing the framework for the collaboration;
    - all outputs from the collaboration are put into the public domain; and
    - an inclusion of a disclaimer stating “The NSO's role is limited to the development of the underlying data source and it is important to note that any analysis, conclusions or recommendations made in this report are the authoring organisation alone”.

15. In the second arrangement, case studies showed a clearer demarcation between the roles of chief statistician and data stewardship. For example, in Mexico, the NSO and the NSS are regulated by a specific law on statistical and geospatial data. The NSO is an autonomous public institution responsible for regulating and coordinating the NSS, and for producing, integrating, and disseminating statistical and geographical information. On the other hand, a separate similarly autonomous public agency is responsible for ensuring compliance with open access to public information and the protection of personal data.

16. In Finland, a representative of the NSO sits on the information management board but there is no clearly defined entity that could be described as a data steward. Instead the functions are the responsibility of many different Government entities. However, the NSO takes an active role in supporting these entities and participating in data stewardship functions.

17. In the final arrangement, the position of a chief data steward or officer (CDO) is defined within the organisation of Government. In some instances, it is a political role and in others it is a career civil servant appointment. The following paragraphs give some examples of the different arrangements for this position given in the case studies where the link between the chief statistician and the chief data steward is often sensitive to political impartiality and independence considerations.

18. In France, the CDO is a political appointment made by the Prime Minister (PM) and reports to the director of the PM’s inter-ministerial task force on open data and includes the functions of a Chief Technology Officer (CTO). The role’s close relation to the PM means it has a unique ability to work across Government. The associated functions of the CDO are distributed among 10-20 members of the task force team.

19. The initial focus of the CDO was on open data and open governments, with some programs to help policymakers make better use of data. While open data continues to be a priority, the scope has now expanded to helping produce and circulate high-quality data. The team also runs a program for public interest entrepreneurs, and a new office looks at data uses based on data science techniques and artificial intelligence. The task force publishes annual reports for the PM, providing a platform to share data policy insights among politicians.

20. In Estonia, the CDO is in the Ministry of Economic Affairs and Communications, serving in the same team as the Chief Information Officer (CIO) and the Chief Technology Officer (CTO). This is a civil service position and not a political appointment. CDO does not have its own staff and the position acts rather like a high-level advisor. The CDO reports to the CIO (vice chancellor), who in turn is overseen by the Ministry’s chancellor. The CDO does not actually produce data, but instead formulates policy, advocates for data use and promotes education around data.

21. In Mexico, a CDO established in 2013 was a cabinet level position housed in the executive branch within the national digital strategy office which reported to the coordinator of this office and the President. The position's main role was the implementation of an open data policy for the federal Government and 2,400 municipalities across 32 states and 300 institutions.

22. The CDO worked closely with the NSO on the technical open data policy despite the sensitivities of collaboration between a political entity and an autonomous organisation. The NSO's involvement helped sustain the project through political cycles and brough invaluable technical expertise to validate project outputs. In turn the CDO had wider reach to coordinate different Government actors with different approaches to data management. Another important aspect of the CDO's work was engagement with external partnerships. In 2018, a shift in Government resulted in the termination of the position. During its existence, the CDO faced several challenges such as;

    - showing the return on investment in opening data for government and private companies is high enough to justify these partnerships;
    - communicating the meaning and characteristics of open data beyond transparency;
    - enacting large-scale reform to have a bigger impact on data policy as the small and impactful projects were not enough to create a catalyst for major changes in data uptake in government;
    - measuring the effectiveness of shifts in open data beyond page views on the portals which is not the best evidence of impacts;
    - the CDO office didn't have a specific budgetary line, so ongoing fundraising from both the government and from external technology funds proved necessary.

23. The city of Los Angeles was one of the first cities to appoint a chief data officer in 2014, specifically to collaborate with city departments on the collection and analysis of data; and manage the city's open data portal. The chief data officer reported to the city controller who manages audit and accountability issues. These reporting line arrangements were a deliberate effort to better inform budgetary allocations with data, as well as to make data investments more efficient across city departments. Since then, the chief data officer role has expanded to be about the creation of innovative, data-driven solutions for use by the city in addressing different public issues, including health, safety, and equity, as well as leading on collaborations with the private sector and other global cities.

National data strategy

24. National data strategies are a foundation to support the role of data steward. They set the ground rules, expectations and aspirations for how data are to be organised and used within Government. Often these data strategies are the responsibility of an authority outside the NSO, but the case studies indicated the NSO can be directly involved in their development and implementation (Germany, Ireland).

25. NSOs did mention the importance of a Code of Practice to help orientate themselves within areas associated with data stewardship. For example, the EU statistics Code of Practice establishes statistical authorities' access to administrative data and NSOs like Finland are involved in the establishment of a range of administrative sources for statistical purposes. In the city of Los Angeles, the data steward is working with partners on a new code of digital ethics and a digital bill of rights.

Open data

26. Another foundation to the data stewardship role is the influence of the open data movement[5]. Several case studies mentioned it as a key impetus for the creation of the data steward role such as in France and the city of Los Angeles; or Germany where an Open Data Institute is being proposed to serve this area. Within Government, the open data movement is usually associated with the creation of open data portals (*data.gov*). These websites and the data within them are usually maintained by Government entities outside of the NSO, but again the case studies indicate a role for NSOs in promoting and practicing open data policies generally. In particular, case study from the NSO of New Zealand mentions their efforts to increase the total available 'open data' resources, but notes opening up more non-sensitive, non-confidential data to the public has continued to be a challenge and the actual use and impact of this increased availability is difficult to measure.

Quality assessment and assurance

27. One of other key aspects the NSO can contribute to a data stewardship approach is its expertise, policy and practices in quality assessment and assurance of data. Colombia and Finland describe in their case studies the development of guidelines and technical standards for the NSS or entire public sector to adopt and adapt for their statistical and data processes. A similar example from Ireland described their lessons learnt from designing a more generalised quality management framework:

    - Should meet all business areas within the NSO but also be scalable to allow it to be rolled out across the entire NSS;
    - Use of a modular framework form, like a design based on the Generic Statistical Business Process Model (GSBPM), meant NSS bodies were able to select those elements relevant to their systems;
    - Implementing organisation's need clear guidance on which data governance elements best fit the statistical and administrative data under their stewardship so that all these data are now managed in a consistent and transparent manner;
    - However, applying a quality management framework to datasets whose purpose is not primarily for analytical/ statistical purposes like those generated from public administration transactions remains a challenge.

Better collaboration

28. Data collected by Government agencies are not always used to their fullest potential nor greatest public benefit. One aspect of this issue is tackling the obstacles to sharing data across Government agencies and ensuring the public maintains trust and confidence in the way data are managed. Trust is not just about privacy and confidentiality controls either; it is also about how these assets or "treasure" as one case study put it are applied in decision making and public discourse. A key enabler to unlocking these public data is collaboration.

[5] Defined by the Open Knowledge International (OKI) as data that can be freely used, re-used and redistributed by anyone – subject only, at most, to the requirement to attribute and share alike.

Many of the case studies referenced using administrative data for statistical purposes and this usually involved some type of collaboration.

29. Three types of collaboration between the NSO and the wider data ecosystem were identified in the case studies which helped foster data stewardship:

    - Coordination – this now usually requires a complex arrangement of committees with representation from all the different Government bodies to coordinate different aspects of data management such as policy, access, open data

    - Facilitation -- there is often a need to facilitate others with the stewardship of their data. As one case study put it, NSOs need to be able to identify “what we can do” from “what we can facilitate others to do”

    - Partnerships -- the NSO needs to partner with other data providers to develop new products or enhance existing ones.

30. This section looks at specific examples or elements of these types of collaboration.

National data coordination bodies

31. Many of the examples given in the case studies focused on their use for official statistics coordination within the NSS, through the use of two coordination bodies types, strategic and technical; where the strategic coordination is focused on data governance and stewardship while the technical coordination is focused on data management such as the execution of processes and practices. A national data coordination body may be in addition or an extension of a national statistical system coordination body or a national statistical coordination body. One example outside of this NSS role was the data ethics advisory group set up by the New Zealand NSO to help maximise the opportunities and benefits from new and emerging uses of data, while responsibly managing potential risk and harms.

Professional networks and secondments

32. The establishment of professional networks was mentioned in several case studies. These could be based on one professional group separated across different organisations in Government or seconded experts from one organisation. For example, both Estonia and France operate a network of high-level ministerial data officers. In Ireland, the NSO seconds statistical staff to other parts of Government to enhance their data analysis capabilities. In the former example, the network is used for coordination but in the latter, it is to facilitate other parts of Government to develop data and statistical capacity. In a sub-national example, the city of Los Angeles organises monthly city-wide collaboratives to help different department leads stay up to date on activities and a data leadership council to bring together department leads with data advocates to discuss strategic data issues.

Data sharing agreements

33. The case studies gave some information on the existence of data sharing agreements between the NSO and other producers within the NSS. In Argentina, NSO access to administrative records is restricted to certain Government bodies, who voluntarily sign

collaboration agreements with the NSO and where compliance with these agreements depends on the continued cooperation of each agency. NSOs reported in their case studies, medical and criminal records continued to be too sensitive to share; and even sub-national data sharing was an issue in one instance. In Finland, several legal acts were revised to enhance the sharing and use of public sector data resulting in several new types of public sector information systems on income, education, health and social protection and local government finances. In Estonia, the NSO gets data from more than 100 administrative sources to create official statistics, but the re-use of these data is limited to the framework of statistical work under which it was collected.

Better methods and capability

34. Data stewardship can be part of transferring methods and capability available in the NSO to other parts of the NSS and even the data ecosystem as a whole or in the reverse, bringing needed methods and capability from outside the NSO, NSS or data ecosystem to within the NSO. This section looks at three areas on this.

Standards and infrastructure

35. It is evident throughout the case studies that several standards developed for official statistics have wider applicability and their dissemination and application across all data actors in the Government can be facilitated through a data stewardship approach. Specific references are made to how the Generic Statistical Business Process Model (GSBPM), Generic Activity Model for Statistical Organisations (GAMSO), the Generic Statistical Information Model (GSIM) and Statistical Data and Metadata eXchange (SDMX) can be applied with in a more generic quality management framework to other data processes and systems of production. Other nationally derived standards are also mentioned.

36. The case studies showed the need to develop infrastructure to facilitate collaboration on data. In Finland, a national data exchange layer was created as a gateway to enable data sharing between different public sector organisations.

Data science, machine learning and artificial intelligence

37. NSOs are keen to acquire further competency in these areas either through partnerships, direct recruitment or other types of skill acquisition. While some of these case studies mentioned this as an area of interest and aspiration, there were few concrete examples from NSO's on how these areas were contributing to data steward roles or opportunities. One exception was the interesting example from the New Zealand NSO on creation of an Algorithm charter to demonstrate a commitment to ensuring public confidence in how government agencies use algorithms.

38. In Estonia, the chief data officer is very active in this area and helped to realize more than 70 different small artificial intelligence projects across Government. Example projects included: satellite images to aid with identifying a range of environmental variables; applying various data to address unemployment needs in training, job search and salary prediction; work with the national library and Ministry of Culture to analyse pictures from a large digital

archive to support researchers, as well as creating metadata for national television; and a project with the state emergency department to improve response times.

More integrated data strategy

39. This area was intended to cover a common approach to data handling. The case studies provided some information on the use and cooperation with geospatial agencies which is very strong in the Latin America region. The NSO of Argentina works closely with the geospatial agency on integration of statistics with geospatial data. While the NSO of Mexico has this responsibility within its own mandate; and has implemented the Statistical and Geographical Production Model (MPEG) – an adaptation of the Generic Statistical Business Process Model (GSBPM) – through which efforts are made to standardize and document statistical and geospatial data production processes, regardless of the thematic domain. Further examples are needed on strategies for how data are integrated across the data ecosystem. Many of the case studies focused on services and products derived from basic statistical infrastructure (i.e. microdata, tabulations, data platforms etc.) where demand remained high; but there was a need to continue to innovate to remain close to users.

Better access and understanding

40. In some case studies, the data stewardship role was an advocacy or spokesperson function without any specific data production or custodian responsibilities. When it is more closely related to the NSO, many of the examples provided were about the traditional NSO role of working with users to provide better access and understanding to statistics rather than data. In this section, four areas are presented where there is potential for NSOs to expand their role in providing better data access and understanding to data in the ecosystem.

Geospatial visualization of the information

41. More case studies are needed to understand how the NSO expertise in the visualisation of the geospatial information can be applied to the wider data ecosystem. For example, the NSO in Ireland built a dashboard to measure national data infrastructure maturity (scale of use of personal, geographic and business identifiers) for Government bodies to identify implementation challenges. The group behind the dashboard also promote and develop solutions to increase take up of these identifiers, such as through the use APIs. A case study from the city of Los Angeles described a geospatial-based platform they created for exploring and visualizing location-based open data, such as information about boundaries, infrastructure, city planning, and health. This also includes a web GIS platform for city employees to facilitate data use and collaboration.

Microdata

42. Several case studies made specific references to the expansion of the provision of microdata. Either to facilitate the use of more sensitive microdata sets across Government derived from national address systems, occupation classes or unique identifier sets for people, businesses and locations such as in France and Ireland; or in the more traditional

sharing of microdata from census and household survey in partnership with an academic facility such as in China.

Data platforms and dashboards

43. The case studies showed a similar need for better ways to present aggregated data through special tabulations, dashboards and data platforms. In France, the office of the chief data steward has responded to the considerable demand for new areas of data relating to Covid-19 with dashboards for politicians and the general public on daily Covid-19 data such locations where mask-wearing is mandatory and collecting, verifying and aggregating data from the field, from the hospitals, different doctors, and other original administrative sources.

44. In the city of Los Angeles, the data steward developed a range of new resources to improve the openness and usability of their data including; an open data portal with 1,000 datasets, a Geohub for exploring and visualising location-based open data, a homelessness data hub, and an open data Covid-19 portal with data on the health and wellness of the city, pandemic related services and information about the reopening of businesses and public spaces.

More integrated data literacy strategy

45. Among the case studies, there were mentions of new approaches to building data literacy. Estonia mentioned funding went towards contracting a university data science master's course, part of a nation-wide effort to promote data education. Similarly, Finland received external funding to develop services for researchers in collaboration with a health and welfare institute.

46. The case study from the city of Los Angeles gave a number of examples of how they lead initiatives to develop a data culture among a broad range of users such as establishment of a new civil society level accreditation as data analyst, community-based training on how to find and use open data resources and working with a university to develop a new curriculum on data science.

How to create an enabling environment at the international level

47. Beyond bringing more attention to the idea of a data stewardship approach, there is a need to bring to more clarity to the definitions and terminology used within this area and apply their translations into different languages. One case study commented that data stewardship in Spanish can be interpreted either as data custody or data management/ administration.

48. There are several other international working groups such as the UN Global Working Group on Big Data for Official Statistics (GWG), the Modernization of official statistics (MOS), the Open Data Working Group and the Expert Group on National Quality Assurance Frameworks (EG-NQAF) whose work touches upon areas related to data stewardship. The working group on data stewardship will ensure to keep appraised on any overlaps and either establish a collaboration on areas of interest or adjust the work programme accordingly to avoid duplication.

Conclusions and next steps

49. The case studies reveal a range of approaches to data stewardship with different levels of involvement from the NSO and provide the main elements to initiate a broader discussion on what it means to adopt a data stewardship approach. However, in order to provide guidance and specific recommendations in this area, as requested by the UN Statistical Commission at its 51st session, further work by a wider group of countries and stakeholders is needed as well as to distribute this information and establish the appropriate terminology in other UN official languages. Beyond a more developed model for data stewardship, it may be possible to develop a series of tools to guide NSO implementation on this topic.

50. The Commission will consider the establishment of a working group composed of NSOs of member states and other stakeholders, with the aim of continuing the stock-taking exercise, compiling best practices, and providing recommendations and guidance on **data stewardship and the role of statistical offices in the new data ecosystem**.

51. The draft Terms of Reference of the Working Group are contained in the report of the HLG-PCCB to the commission.[6]

[6] E/CN.3/2021/4.

Annex I: scoping methodology

i. This paper proposes a *theory of change* methodology[7] is applied to the approaches to data stewardship work stream to help identify the key drivers of change needed to adopt or support the role of a data steward in any given data ecosystem. This method's starting point is to define a desired impact statement or goal to be attained and then map backwards the necessary pre-conditions and foundations such as legal, technical and institutional mandates. Under this methodology we define the broad goal of a data steward is to:

Improve the use of data in society

ii. Following this method, a set of outcomes are defined which will lead to this goal. Since improved use is determined by fitness for purpose, the desired outcomes are informed by the common quality dimensions[8] used in official statistics, which are derived from the Fundamental Principles of Official Statistics (FPOS)[9]. Four outcomes are proposed as follows:

Better governance	More oversight and better governance by establishing a strong independent role of the data and statistical authority to provide an expert opinion of data produced within Government and across society to ensure trust, protect data rights and privacy, and counter misinformation and data misuse.
Better collaboration	More collaboration between producers and users within Government and across other communities (i.e. private sector, academia etc.) and improved coordination by brokering new partnerships to ensure data of public interest[10] that is relevant, interoperable, comparable, and meets needs.
Better methods	More transparency to ensure openness to a range of new data methods, sources and technologies from different parts of the statistical and data systems, public and private, and mitigating latency in information systems
Better access	More consultation to bring in the viewpoint of the user and

[7] See https://en.wikipedia.org/wiki/Theory_of_change for further references on the methodology.
[8] See https://unstats.un.org/unsd/methodology/dataquality/references/1902216-UNNQAFManual-WEB.pdf
9 See https://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf.
[10] Public interest is a common concern among citizens in the management and affairs of local, state, and national government. It does not mean mere curiosity but is a broad term that refers to the body politic and the public weal.

enable users to work with data more easily and target population groups who maybe have more challenges to using data fully in their lives

iii. These outcomes need to be supported by outputs which bring about these changes which in turn are supported by activities to achieve each output and the inputs needed to accomplish the activity. Since there is no common view on what data stewardship requires, this paper proposes some to initiate a discussion about what outputs, activities and inputs are needed to bring about these outcomes (see the table at the end of this section). Over the course of the stock-take, it is hoped a better understanding emerges of how NSOs can position themselves within the National Data Ecosystem and the kind of changes a data stewardship approach could usher in. The stock-take will acknowledge and incorporate different starting points, institutional environments and national/ regional set-ups to be relevant to most settings and situations.

Working mechanism: development of case studies

iv. Case studies on aspects of a data stewardship model (as defined in Table) will be the main source of information to inform the stock-take. The type of "data" referred to in the impact statement will be informed by the case studies collected, but "data of public interest" (see earlier definition) are a suggested starting point.

v. Different forms of data stewardship being followed in data ecosystems will be used. In addition, it would be useful to look at examples by different types of organisational and data system circumstance like level of capacity[11], the political context, institutional set-up (i.e. centralised or decentralised statistical system at the national level or even localised or city statistical systems at the sub-national level) and those NSOs within countries with special circumstances (i.e. small islands, least developed, landlocked etc.)

vi. The case studies will provide a better understanding of the roles and activities of a data steward and how they fit into different contexts (i.e. capacity, institutional, organisation). This will then give the basis for an initial set of options or drivers of change to be considered for further discussion by the global statistical and data community. The findings as well as any gaps in knowledge will be presented at the 52nd session of the UN Statistical Commission as part of the official report of the HLG-PCCB as well as presented in a more detailed background report and will provide the basis for further work by the commission and the broader data community

[11] http://datatopics.worldbank.org/statisticalcapacity/

Scope of the “data stewardship” stock-take

Outcome	Outputs	Activities	Inputs / needs
Better governance	Role of Chief data officer / national data steward / another institutional role	Coordination within national context and institutional setup to understand the best way to establish and maintain leadership and direction	- What is the scope of the role? - Implications for the NSOs changing functions (i.e. capacity and mandate)? - How is the role complementary to the chief statistician? - How does it fit in an existing NSO’s organisational structure (as new position or part of existing one)? - What is the role in setting ethical standards and safeguarding privacy of data to ensure trust in data? - What are the key elements to assess the NSO baseline related to data stewardship implementation approach? - What are the key elements to consider a shift from statistics producers to service providers? - How does the NSO's autonomy and independence relate to a role of stewardship?

Better governance	National data strategy		- What is the purpose of these strategies and how do they differ from code of practice and other documents released by the NSO, including the UN Fundamental Principles of Official Statistics? - How does professional independence and accountability of NSOs impact on data stewardship role? - Should a Fundamental Principles for data of public interest be created at the international level to reflect norm setting
	Open data	Adoption of open data practices and expansion of open data policies to more sets of data	- How best to use existing assessment tools (i.e. ODIN)? -

Better governance	Quality assessment and assurance	Build out skills within the organisation(s) such as data officers (skills investment)	- How can professionals within the NSO be used to enable data stewardship? - Can existing frameworks be adapted easily to indicate quality of non-official statistics and non-Government data? - How can official statistics methods and quality frameworks become a quality certification to be applied to non-official and non-governmental statistics? - How do NSO policies and practices on these areas change with a data stewardship approach?
Better collaboration	National data coordination bodies	Co-ordination role of the NSS on data sharing	Are data coordination bodies part of or separate to NSS and NSO coordination bodies? How do they relate?
	Professional networks and secondments	Facilitation role of the NSO	What role, if any, should the NSO play in professional networks and seconds in the national data ecosystem?

Better collaboration	Data sharing agreements, which safeguard data security and privacy	Partnership between the NSO, NSS and wider data ecosystem	- What are the challenges of data sharing agreements with different partners (i.e. Govt, private sector, legal capacity challenges, academia etc.) in a data ecosystem? Can lessons from NSOs be applied in the wider data ecosystem? - How can countries improve their legislation to adopt a data stewardship approach (i.e. access to administrative public or private data, and big data) - What does a standard agreement format look like to be applied to the new data ecosystem?
Better methods	Standards and infrastructure		- Does data stewardship require investment in any new technology? - What opportunities are there for public-private partnerships? - How does policies and techniques of security and privacy under official statistics apply to the wider data system? What standards and infrastructure from the statistical system be adopted in the wider data ecosystem?

Better methods	Data science, machine learning and artificial intelligence		- How does incorporating data science within the NSO assist with collaboration and partnerships in a data ecosystem?
	More integrated data strategy	Creation of a common approach to data handling	- How can data capitalize from linkages or calibration with statistical sources? - How do NSO engagement on data change with a data stewardship approach and encourage greater responsibility for innovation? - How can data insights from NSOs and new data actors complement each other in the new data ecosystem? - Are there alternatives to obtain insights from data without seeing the data owned by a third party, segregated by jurisdictional boundaries? - How can the international statistical community achieve agreements or partnerships with the big tech monopolies to have access to data and new data analysis techniques? - Should the NSO guarantee the free access to public data that private key stakeholders have? - What kind of technological improvements should the NSO propose to the NSS in order to handle properly this data sources?

Better access	More integrated data literacy strategy	Learning (including data science)	- How does data stewardship expand the NSO role for informing users and training its own staff, especially on subjects like data science? - How do programmes reach youth and other groups who need more support to use data fully? - What is the return on investment of the NSO in the new data ecosystem regarding statistical culture? - How should the NSO and academy agreements look like in order to speed up the data scientist formation in the NSO and other government sectors?
	Geospatial visualization of the information	Integration of geospatial systems and information with the statistical production and dissemination systems	- How to guarantee that more data is geo-referenced? - What does a proper standardization process of geospatial information look like?
	Metadata	Application of metadata standards to more data in the data ecosystem	- How to approach this task? - How to create acceptance and buy-in, such as for common standards like Data Documentation Initiative (DDI) and Data Catalogue Vocabulary (DCAT)?

Better access	Microdata	Expansion of microdata facilities to all government data holdings	- Adaptation of the work being undertaken by the Inter-secretariat Working Group on Household Surveys - what role can NSOs play in microdata access to government data holdings? Are an NSO's microdata access facilities a tool for all government data custodians? - What legal and privacy roles and responsibilities does an NSO have for non-statistical data holdings?
	Data platforms and dashboards	Facilitate the expansion of interoperability and data platforms among other producers in the data ecosystem	- How can interoperability be expanded to different data custodians?