United Nations ECE/CES/2023/2

Economic and Social Council

Distr.: General
13 April 2023

Original: English

Economic Commission for Europe

Conference of European Statisticians

Seventy-first plenary session

Geneva, 22–23 June 2023

Item 5 (b) of the provisional agenda

Reports, guidelines and recommendations prepared under the umbrella of the Conference

Data stewardship

Data Stewardship and the role of national statistical offices in the new data ecosystem

Prepared by the Conference of European Statisticians' Task Force

Summary

This document is an extract from "Data stewardship and the role of national statistical offices (NSOs) in the new data ecosystem" prepared by a dedicated Task Force composed of Austria, Belgium, Canada, Denmark, Estonia (Chair), Finland, Hungary, Ireland, Kazakhstan, Mexico, New Zealand, Netherlands (the Kingdom of), Poland, Switzerland, the United Kingdom, Bank of International Settlements (BIS), Eurostat, the Organisation for Economic Co-operation and Development (OECD) and UNECE. The members represented NSOs and other government agencies, central banks and geospatial agencies.

The first draft of the report was consulted electronically with all CES members in spring 2022. The feedback provided support and many constructive comments how to improve the document. The Task Force has updated the document, taking into account the comments and the latest developments and developed also two supplementary documents: (1) a summary of the report, which can be used as a self-standing document to explain NSOs' role in data stewardship and (2) a document with examples of how NSOs in different countries are pursuing their role as data stewards. The CES Bureau reviewed the draft report in February 2023 and asked the Secretariat to send it for electronic consultation to all CES members before the 2023 CES plenary session.

The current document is a short version of the report prepared for translation purposes. The full text, which includes also chapters on data governance models in the public sector, data stewardship maturity assessment model, communication about data stewardship, conclusions and annexes, and above-mentioned supplementary documents were sent to all CES members for consultation in April 2023 and are available on the Conference webpage at https://unece.org/statistics/events/CES2023. Summary of feedback received during the consultation will be provided in document ECE/CES/2023/3/Add.1.

Subject to a positive outcome of the consultation, CES will be invited to endorse the document.

I. Executive summary

1. The rapidly expanding and evolving landscape of data production and sharing is an opportunity for National Statistical Offices (NSOs) to exercise their expertise in the larger statistical as well as national data community. Indeed, because of this environment, NSOs must expand their role into data stewardship in order to fully carry out their mandate. As the transformation is multifaceted, there are different interpretations of how the NSO role can and should change, and what data stewardship means in this context.

2. In this report, the United Nations Economic Commission of Europe (UNECE) Task Force (created by the Conference of European Statisticians in 2021) has examined the issue of data stewardship through the lens of NSOs. Data stewardship is viewed as operating in service of, rather than in control of the data ecosystem and reflects an overall paradigm shift or evolution of the framework that guides how NSOs operate, moving from the production of statistics to the provision of data and data-related services.

3. The work of this Task Force has focused on defining and explaining the foundational concepts and frameworks relevant to data stewardship. More concrete guidance and recommendations could be developed at a later stage, as a follow-up to this report. The Task Force has considered further steps to support NSOs in implementing their stewardship role, which are summarised in Section IB ‘Future work’.

4. Ultimately, the report aims to serve as a guide to statistical offices regardless of data or structural maturity. It will help offices decide whether they should extend their role as government data stewards, as well as provide inspiration and ideas to those who are already performing this function. Public data holders will also gain a deeper understanding of the advantages and services provided by NSOs in the future data ecosystem. Below is a high-level summary of the report as a whole, followed by a distillation of the key messages that the report communicates specifically with respect to data stewardship.

A. Report highlights

5. The development of complex data supply chains, new technological developments, and increasing importance of infrastructure for data sharing and access requires monitoring and maintenance. This has put a renewed focus on public trust and subsequently, highlighted the importance of data stewardship and governance (Chapter 2 in the full report). NSOs can offer competencies and capabilities to contribute to better governance of data in the public sector and beyond, while bearing in mind the legal, organisational and technical aspects that shape NSOs’ role in this area.

6. Common definitions of data governance and data stewardship are necessary, and can be leveraged in the context of official statistics and across the public sector (Chapter 3 in the full report). Data stewardship is a particularly important concept in current data operating environments, especially in the way it promotes a holistic perspective on the management of data though clear governance accountabilities that emphasizes intergenerational guardianship and public good.

7. Data stewardship is enabled through clear and authoritative data governance. To organise data governance activities in the public sector, three types of enabling governance models include centralised (top-down), distributed (bottom-up), and federated (hybrid) data governance (Chapter 4 in the full report). The centralised model is characterised by consolidated governance authority, and the distributed model by independent governance. The federated model is a hybrid model with a central authority that enables input from agencies, and is the model best suited for an all-of-government approach to data stewardship. Selection and implementation of a data governance model will depend heavily on national contexts.

8. The successful delivery of data stewardship relies heavily upon people’s understanding of stewardship, and especially how it relates to their roles and functions – both within an agency and at the system-wide level (Chapter 5 in the full report). The roles and corresponding competencies of a data steward can be explored through examples and case-

studies from different countries, each with their own unique context and particular lesson-learnt.

9. Measuring data stewardship performance is relatively new territory, so a data stewardship maturity assessment model is proposed. A data stewardship maturity model can embed important ideas and values into the understanding of what constitutes sound data stewardship and, by extension, into organisational structure and practice, to steer agencies in the right direction in the digital data ecosystem (Chapter 6 in the full report). A data stewardship maturity model can help clarify the meaning of data governance and highlight the ways it is meant to operate in conjunction with, and help deliver, data stewardship. One such possible maturity assessment model is presented, based on an example implemented in New Zealand.

10. Though data stewardship can represent a challenging concept to understand, the importance of data stewardship and NSOs' changing roles in performing it can be effectively communicated using enabling products, services, and partnerships (Chapter 7 in the full report).

11. This commitment to data stewardship and addressing the changing data landscape can result in new partnerships and data sources which decrease the response burden and increase the timeliness, quality, interoperability, reusability, and fitness-for-purpose of data, contributing to better insights and services to citizens (Chapter 8 in the full report).

B. Key messages about data stewardship

12. The concept of data stewardship is fundamental to data best-practice (especially in an evolving ecosystem) and to the future of the work of NSOs. To provide clarity and consistency in this subject, the Task Force has developed a cohesive definition:

> *"Data stewardship means ensuring the ethical and responsible creation, collection, management, use and reuse of data so that they are used for public good and benefit the full community of data users."*

13. The report will convey the following key messages about data stewardship:

14. Data stewardship is necessary to maximize the value of data assets.

- Stewardship enables data policy implementation and the treatment of data as a strategic asset.
- Stewardship promotes sharing and reuse of data assets subject to the right ethical and cultural conditions, thereby maximizing their value .
- Stewardship promotes open data, interoperability and ethical and culturally appropriate use of data, enhancing the timeliness and efficacy of decision making and enabling data-driven public service delivery.

15. Data stewardship has two main scopes: intra-agency and system wide.

- Internal data stewardship can help consistently promote best practices in working with data within an agency, including improving data quality and metadata-based production .
- System-wide data stewardship requires cooperation and coordination across agencies which goes beyond various data domains. Achieving the desired availability and interoperability of data is therefore a multi-party effort.
- There is no 'one-size-fits-all' approach to data stewardship and the needs of individual NSOs must be considered.

16. NSOs have inherent and unique expertise to lead data stewardship in the National Statistical System (NSS) and to take on data stewardship responsibilities across the national data ecosystem.

- In their role of producing quality statistics, NSOs have always been engaged in data stewardship. This unique experience puts NSOs in a position to leverage its

experience and functional expertise to take on new tasks and expanded roles related to data stewardship.

- The data stewardship role of the NSO can cover responsibilities in system-wide data stewardship related to public sector data management, depending on the national context.

17. Data stewardship improves data quality and relevance of NSOs and the National Statistical System.

- Data stewardship can support the delivery of high-quality data in all countries. Data stewardship brings added value to NSOs and the NSS, factoring for the national context affecting its implementation.

18. Data stewardship strengthens public trust in official statistics and in data management across the public sector.

- Provision of data-driven public goods and services requires sound data governance, public trust and data reliability.
- Data stewardship supports the Fundamental Principles of Official Statistics which enable privacy, high quality and public trust.

19. Data stewardship needs governance, capabilities and resources for successful implementation.

- Data stewardship goals should be embedded in data governance - a system of decision rights and accountabilities for the management of data, and the resulting laws, regulations, policies and frameworks that provide enforcement.
- Data stewardship requires co-ordinated investments and capabilities, in transparent dialogue with the public and the budget decision makers to ensure progress and needed investments.
- A maturity model can be used for self-evaluating the gaps between a current state and desired end-state of data stewardship. This allows the definition of appropriate governance structures and planning investments.

I. Introduction

A. Background

1. Statistical offices operate in an increasingly interconnected, dynamic and digitized world, which brings with it new demands and new opportunities. Data has become a vital strategic asset and as such, many organizations outside official statistics create statistical output using new data sources, tools and techniques. This has also changed stakeholders' expectations for more open and detailed data that is both quickly and easily accessible. The COVID pandemic has further tested official statistics' ability to harness quality and timely data in a responsible way, accelerating a modernisation in how NSOs, public sector agencies and private organisations deal with data.

2. The data and information potential from leveraging partnerships and linking data from multiple sources, raises questions related to ethics, privacy protection, data security, accessibility, harmonisation of concepts, formats and technologies used, knowledge of data available, and how data will be managed and governed in the long-term. To respond to these questions, this report will define the term 'data stewardship' and explain its functions and competencies to enable access to and re-use of data for public benefit in a systematic, sustainable and responsible way (Verhulst, 2021).

3. To move towards a common understanding of the term data stewardship, the Conference of European Statisticians (CES) set up a Task Force in 2021 to help bring clarity to this area and facilitate the work of NSOs in adapting their role to the changing situation. The following countries and organizations are members of the Task Force: Austria, Belgium, Canada, Denmark, Estonia (Chair), Finland, Hungary, Ireland, Kazakhstan, Mexico, Netherlands, New Zealand, Poland, Switzerland, UK, Bank for International Settlements (BIS), Eurostat and Organisation for Economic Co-operation and Development (OECD). The Task Force members represent statistical offices and other public institutions, and the UNECE is providing the Secretariat.

4. A draft version of the document was circulated to the countries and organizations participating in the work of the Conference of European Statisticians (CES) in spring 2022 for consultation. The present draft document incorporated the vast and varied feedback, and latest development. The CES Bureau reviewed it is February 2023, and asked the Secretariat to send it for wide consultation. Subject to a positive outcome of the consultation, the CES 2023 plenary session will be invited to endorse the document.

5. The report aims to serve as a guide to statistical offices regardless of data or structural maturity. It will help offices decide whether they should extend their role as government data stewards, as well as provide inspiration and ideas to those who are already performing this function. Public data holders will also gain a deeper understanding of the advantages and services provided by NSOs in the future data ecosystem.

6. The role of national statistical offices in the data ecosystem has been widely discussed during recent years. However, the discussions vary considerably in focus and there is no common understanding on what data stewardship is, or what the roles of NSOs in this evolving context can (or should) be. Relevant questions include:

- What is data stewardship? What is its relationship to data governance?
- What role might NSOs play in data stewardship in the public sector, especially in cooperating with other components of the NSS, or more widely in a country's data ecosystem?
- How will NSOs respond to competition from other organisations and what strategic partnerships are needed?
- What criteria is used by NSOs when deciding how to manage their position as data stewardship subject matter experts, for the public sector? What does their leadership look like?

B. Future work

7. When the Task Force was set up in February 2021, it was noted that the report could serve as a basis for developing further guidance and recommendations on the role of NSOs in data stewardship. In international discussions on this topic, countries have asked for practical guidance on how to implement the Data Stewardship role.

8. To provide support for practical implementation, the Task Force has considered the following as potential next steps:

- Develop Basic Principles of Data Stewardship by monitoring and analysing implementation of data policies and data governance initiatives by NSOs
- Provide more specific guidance and/or recommendations and tools for implementing data stewardship
- Develop a generic roadmap for how to increase NSOs leadership role in data stewardship (taking as a starting point a paper by Estonia et al., presented at the 2020 CES plenary session)
- Identify a core set of responsibilities of NSOs as data stewards in national data ecosystems
- Compile a knowledge base of best practices and examples of successful data stewardship implementation, including more detailed examples of how data stewardship is implemented and the data services provided in practice
- Collect examples of maturity models and other assessments of data stewardship, and on this basis, develop and test a generic maturity model for the assessment of NSO data stewardship
- Prepare guidance materials for different audiences or focusing on data stewardship in specific areas - administrative data, open data, privately held data, geospatial data, governance models, etc.
- Provide a forum for exchange of experience in implementing data stewardship
- Maintain the glossary (consistent with the work of the Unites Nations Statistics Division’s global Working Group on Data Stewardship) and facilitate the translation of key terms into other languages.

9. In undertaking any follow-up activities, it is important to consider other global and regional work on this topic, including the above-mentioned global Working Group on Data Stewardship (WGDS) with its five work streams, and the work in the Economic Commission of Latin America and the Caribbean (ECLAC) and The Economic and Social Commission for Asia and the Pacific (ESCAP) regions. Any follow-up proposals will be submitted to the CES Bureau in October 2023.

II. National statistical offices and data stewardship

A. Context

10. With the exponential growth in data generated worldwide and the arrival of technologies such as artificial intelligence (AI) and edge computing, public authorities worldwide are increasingly recognizing the importance of data. In this context, it becomes vitally important to set the data stewardship and governance foundations that can enable data access and sharing, while protecting the privacy of individuals and ensuring that the data are used ethically.

11. The use of electronic devices through communication networks has allowed companies to build up rich data pools, for example, in online commerce and entertainment platforms. This consumer data, combined with other – data (whether public or private) offer enormous potential to discover patterns, infer indicators, develop business models, or make decisions for the common good. Access to these data pools is currently restricted due to

intense market competition, regulations on privacy and competition, as well as the implications associated with the costs of obtaining and sharing information. The grassroots development of joint public and private sector initiatives has made it possible to explore the benefits and challenges of sharing information held by private businesses. However, this case-by-case approach has limited the potential to produce regular statistical information for public use.

12. The digitalisation of society and economy has placed data access and sharing at the core of innovation and public trust. The use of administrative registers and experimental techniques allow data to be obtained from other sources and for different purposes. These advances offer the opportunity to supplement the statistics that NSOs have traditionally been producing, thereby reducing administrative cost, and response burdens.

13. These developments offer both opportunities and challenges. Government agencies and other public institutions have their own information systems and data collections with rich histories, which have led to countless data holdings that can be enriched when supplemented with alternative data. In addition to reducing burden, there is huge potential for new types of data services, more timely and granular data, new insights by linking data from different sectors, sources and topics. However, there are also huge risks that need to be mitigated. Ungoverned data could be used unethically, increased digital capacity comes with an increased need for digital security, the 'digital divide' could become an 'information divide', and of course, NSOs are challenged to remain vigilant and proactive in the protection of privacy as these developments occur. Further, in order to benefit from these developments, public agencies will have to address issues around a lack of data accessibility, interoperability and standardization; data duplication and redundancy; and the costs of linking and sharing data.

14. To overcome these risks requires governmental policies with farther foresight, to enable the focus on things like strategy, culture, ethics, roles, and the capabilities of people that can support an effective data ecosystem. The concept of data governance requires common rules and standards to make interoperability possible. For this to work at an all-of-government level, there must be coordination, accepted norms and adopted standards. Fortunately, a strong culture of coordinated collaboration has been implemented within the NSS, not least in the context of the application of the Fundamental Principles.

15. Data governance establishes and makes accessible the set of norms and standards associated with data. Agreed upon data governance norms will provide clarity and assurance in the way data is overseen and managed, and will support data availability, quality, security, usability and integrity. While data governance has a strong emphasis on technical competencies, data stewardship focuses equally on the foundational concepts, frameworks, and best-practices to support the growing maturity of data policy.

16. Data stewardship implements the policies, standards and principles outlined by data governance. It encompasses functions and competencies to enable access to, and re-use of, data for public benefit in a systematic, sustainable, and responsible way. Data stewardship is operating in service of - rather than in control of - the data ecosystem.

17. National governments are taking steps to reduce policy siloes and make public sector data strategies, projects and initiatives more coherent (OECD, 2019). Countries such as the United States, the United Kingdom, the Netherlands and Germany have issued National Data Strategies to cover different elements of the data ecosystem ranging from data access and sharing, open data, data for AI, and data ethics in the public sector – all under a single policy instrument. At the international level, there are ongoing efforts to promote data sharing and data access, such as the G20 Data Gaps Initiative (DGI). Its new phase, which launched in 2022, covers among other recommendations, access to private and administrative sources of information and data sharing (FSB Secretariat and IMF Staff, 2022).

18. In other cases, data-related policy instruments are made a part of broader digitalisation strategies. Indeed, data from the 2019 OECD Digital Government Index show that "only 12% of countries have a single dedicated data policy (or strategy), while 82% embed data as part of broader related policies (e.g. digital government or open data)" (OECD, 2020).

19. In terms of institutional governance and the creation of enshrined leadership roles, some countries have opted to define specific Chief Data Officer (CDO) positions for the whole-of-government (e.g. the appointment of a Chief Data Officer for the Government of Canada in 2022), while others have opted for creating bodies such as data councils where data leadership roles at the institutional level discuss priorities and agree on their coordinated implementation (e.g. the U.S. Federal CDO Council).

B. What National Statistical Offices have to offer

20. National Statistical Offices are mandated to produce and disseminate relevant data and information in coordination with other parts of the NSS, in order to understand the evolution of society, the environment and the economy. This mandate comes from the United Nations' Fundamental Principles of Official Statistics and is enshrined in each country's statistical legislation. For decades, NSOs have developed methods and techniques to produce the insights necessary to meet this mandate. Concurrently, they consistently and collaboratively promote a culture of using data and information strategically in making decisions for the public good.

21. The United Nations' Fundamental Principles of Official Statistics provide internationally endorsed principles for how NSOs can best provide their services to citizens (United Nations 2014). To summarize, the Fundamental Principles state that official statistics must be:

- Practical and impartial
- Accurate, trustworthy and ethical
- Scientific and standardized
- Quality, timely, and with as low a response burden as possible
- Confidential and appropriately used
- Transparent and made public.

22. The Fundamental Principles also state that NSOs and other statistical agencies can draw data from a variety of sources, coordinate and cooperate with other organisations (nationally and internationally), and contribute to the improvement of statistical systems (United Nations, 2014).

23. Developed and adopted by the Conference of European Statisticians, the United Nations Fundamental Principles of Official Statistics (FPOS) were adopted in 1992 at the ministerial level by the Economic Commission of Europe (UNECE) (United Nations, 2021a). These Fundamental Principles were then reaffirmed and endorsed by the UN Statistical Commission at the global level (1994) and the UN General Assembly (1994) (United Nations, 2021a). The preamble of FPOS highlights the importance of official statistics for national and global development, emphasizing the vital role that NSOs play in providing high-quality, official statistical data, and the positive effect that this has on policy decision-making and the overall development of countries (United Nations, 2014). The link between Fundamental Principles and data stewardship is explained in more detail in Annex 2.

24. NSOs have attained a certain reputation as a result of the technical strength, reliability, and timeliness they have displayed during the diligent management of the information they collect. The track record of NSOs is a testimony to their experience and competence in working with data, including confidentiality and privacy protection, data ethics, ensuring comparability, use of standards and classifications, communicating insights, supporting and growing data literacy, etc. While NSOs are seen primarily as producers of statistics, their expertise outlined above and their experiences in their role in coordinating activities within the NSS could be leveraged as an asset for the public sector and beyond. In turn, NSOs can also benefit from other public sector agencies' experience in tackling these issues in their specific domains.

25. National Statistical Offices face the challenge – and the opportunity – to place themselves as key players in the data ecosystem given their well-grounded expertise in data management, access and sharing practices in the public sector, as well as their key role as producers of statistics and indicators. Figure 1 provides a summary of capabilities that NSOs have which give them a good basis to fulfil the data stewardship role.

Figure 1.
Capabilities of NSOs to take on responsibilities in the data stewardship landscape

26. The following frameworks and expertise may be leveraged by NSOs to provide data, information, insights and visualizations to respond to any emerging need of agencies, citizens, businesses and other institutions:

- Legislative frameworks and legal authority to collect, process and disseminate data and information (based on the Statistical Legislation)
- Well established data quality frameworks including standardised metadata, reference data and master data
- Long history of data protection and leading-edge methods to protect privacy and confidentiality
- Modern statistical methods for data analysis, data integration and linkage, data categorization via metadata and data visualization
- Growing data science expertise grounded in statistical and mathematical theory, expertise in modelling and forecasting
- Expertise in data ethics and growing experience in the ethical use of artificial intelligence and machine learning to ensure methods are without bias
- Trust frameworks and development of new frameworks such as proportionality and necessity
- Storytelling and data visualization expertise from data to insights to policy to services for better outcomes
- Well established culture of coordination within the NSS, and access to resources and networks with other NSS members, and a long history of data partnerships and collaboration both domestically and internationally
- Collaborative workspaces (protected cloud, data collaboratives, data trusts), experience in data-access services for on-line access to data by trusted and certified external users
- Statistical and data standards and classification systems to give data meaning and drive interoperability
- Data collection expertise (censuses, different types of surveys, use of administrative registers, web-scraping, etc.)
- Well established data dissemination vehicles including web presence, data hubs, data hosting services

- Emerging entrepreneurial spirit and culture of continuous learning and development of innovation ecosystems.

27. Beyond sound data management and standards, stewardship also requires public acceptability and trust. Social acceptance and public engagement are directly linked to privacy, confidentially, security and transparency, and they are necessary for NSOs to be able to perform their duties and fulfil their mandates. NSO's regular commitment or adherence to the principles of open data, transparent processes, ensuring confidentiality and security, and communicating the value of citizens' sharing their data, all constitute further evidence of their expertise and opportune placement within statistical systems.

28. Data stewardship is an extension of the role that NSOs have played in coordinating the national statistical system, where the key objective is to ensure good coordination between statistical agencies within countries (Fundamental Principles of Official Statistics 8) (United Nations, 2014). The NSOs role has the potential to go deeper into the statistical system to ensure interoperability, and to go wider, to the whole data ecosystem.

29. However, the implementation of this leadership role in data stewardship is highly context-contingent and will vary depending on national factors such as legal frameworks and NSOs' mandates, maturity and use of digital technology, the nature of the data ecosystems, the extent of cooperation achieved within the NSS and public acceptance. Hence, the extent to which NSOs can provide leadership in data governance and stewardship will vary from country to country.

30. NSOs can position themselves in the data ecosystem in different ways ranging from minimal change to the NSO role, to operating as a public sector data steward with the requisite governance authority. Internationally, NSOs are all in different places with respect to their maturation journeys, with some already having established cross-government leadership roles. There is no "one-size-fits-all" solution that can be applied to all countries. In moving towards data stewardship, it is important to consider the jurisdictional and legal context and recognize that NSOs will move forward at their own pace, and taking on a stewardship role that fits their purpose and environment.

C. What is to be gained, and what is at stake?

31. Change sometimes involves risk, and regardless of the decision that an NSO makes regarding data stewardship, there are risks:

- If NSOs fail to participate (at all or in a leading capacity) in the process of data ecosystem transformation, not only do they risk losing relevance but there is a high risk of setting up parallel data stewardship structures which will increase instead of reducing burden and cost. The outcome could be further fragmentation of public data holdings with little or no compatibility.
- On the other hand, there is also the risk of taking on too ambitious of a data stewardship role, one which may compromise NSO's core operations and lead to reputational losses or decreased data quality.
- Further, the aspects of cyber security cannot be downplayed. Even though NSOs already apply the highest IT security standards, as data stewards, NSOs are more likely to become potential targets for cyber-attacks and need to take action to be prepared for this.

32. When successfully carried out, the data steward role makes it possible for NSOs to deliver an important and much needed service to the public while also increasing the positive effect that they can have on data stewardship and quality across the public sector. It will demonstrate that NSOs can fulfil their core responsibility, while at the same time increasing the quality and quantity of data and official statistics. NSOs will benefit by having secured an important role into the future that will provide them with:

- Access to more data sources allowing to produce more timely, relevant and disaggregated data

- Increased possibilities for integrating data from different sources (especially if there are common identifiers).

33. Sound data stewardship has societal benefit and contributes to the public good because it enables ethical operation, which creates the trust, social acceptability and public support necessary for statistical work. It minimises data misuse and enables reuse, allowing statisticians to access data already in the ecosystem. It also facilitates data- and knowledge-sharing and the use of new, complementary data sources. All of this either directly or indirectly improves public trust and increases engagement, by saving time and money, decreasing response burden, increasing data value, and better enabling the communication of that value to citizens.

D. Enabling environment

34. Several enabling factors and trends are already in place in most countries that facilitate NSOs taking a proactive data stewardship role. Three of these factors are described below: 1) legal and regulatory frameworks; 2) technological developments; and 3) organisational interoperability.

1. Legal and regulatory frameworks

35. The production of official statistics in countries and the role of NSOs is based on statistical legislation. This legislation looks different country to country and may outline direct or indirect mandates or role designations to the NSO for data stewardship - ranging from full subject-matter leadership to a particular function as a part of a wider NSS framework. However, the absence of a legal mandate should not stop NSO from exploring the benefits of taking on a stewardship role in the data ecosystem.

36. In general, legal regulations concerning data and information are rather new compared with many other regulated issues. As they reflect the digital and data landscape, they have naturally been developed according to individual countries' contexts. Only countries in the European Union have harmonized legislation on issues like data protection. In addition to being responded to differently country to country, many data issues are constantly evolving, and as such, are in a perpetual drafting process at both the national and supra-national levels. NSOs are well positioned to play an active and leading role here, providing expertise as well as consistency and ultimately efficiency to the drafting process.

37. When constructing a legal environment to enable data stewardship, it must address the implications of reuse and sharing throughout the data life cycle – from collection to dissemination. It will be necessary to establish the responsibilities of those who guard and manage the information (data stewards) as well as the legal and administrative basis (agreements, other policy instruments, etc.) under which this is done, and the mechanisms (federated, centralized, platforms, trusts, etc.) by which it will be achieved.

38. The adaptation of the legal framework must be consistent with national data strategies and compliment a flexible system that allows operational adjustments without the need to frequently modify the laws. That is, it must recognize the continuous expansion of the borders of the data ecosystem, resulting from technological development and the evolution of economic systems and public service models; and it must establish the obligations and rights for all actors in the ecosystem.

39. The rules associated with the operationalization of information reuse mechanisms must be administrative in nature, ideally designed by consensus of the actors involved (standards, models, etc.). In this way, if it is necessary to adjust them in the presence of new risks or technological changes, their modification can be organized without the need for legislative approval.

40. The legal frameworks that govern NSOs' data stewardship activities, must support third parties to carry on research and data-driven projects. It also enables the identification of new third-party sources, to create and calculate statistical indicators. The technical and legal challenge is then to ensure that this information is regularly available, while

guaranteeing that its reuse does not legally compromise any of the parties that participate in the generation of the data.

2. Technological developments

41. The technological environment is constantly developing and organisations across jurisdictions are facing the challenge of how to benefit from these enabling technologies. The maturity of technology implementation differs widely from country to country, but regardless, there are several data stewardship-related requirements that NSOs should prioritize: (1) achieving the appropriate level of digitalization; (2) effective registers and use of digital services (i.e., using administrative data sources; (3) governed data architecture, and master and reference data management; (4) data openness and protection, data privacy and cybersecurity; (5) harnessing artificial intelligence (AI) technologies, machine learning and big data; (6) nation-wide information systems and economic sectors or domain data spaces.

3. Organisational interoperability

42. The extent to which NSOs can perform a data stewardship role, and the successful implementation of this role, depends on the maturity of the data ecosystem in the whole country. This maturity is impacted by legal and technological enablers, but also organisational enablers such as existing coordination arrangements, public service departmental data strategies, national data strategies, or other frameworks or systems that provide a course of action for managing and leveraging data as a strategic asset. When such strategies are developed, NSOs should be actively involved in the discussions based on their expertise and experience coordinating activities within the NSS, managing data collections, advancing interoperability, supporting and enabling standardisation, ensuring privacy and confidentiality, maintaining data security and growing data literacy.

43. Data stewardship is not only for developed countries where the government data systems are already established. For example, an NSO should seek to influence the establishment of administrative registers to ensure that they are appropriate to use for producing statistics. Proposing common approaches for data architecture, standards and classifications before the public data holdings are set up, will allow to save resources and get out more value from data in the long run.

III. Defining data stewardship and data governance

A. Data stewardship and related terms

44. Data stewardship operates in a context of a data ecosystem; therefore, we start from defining this term. The data (including statistical data), along with the data subjects, a broad range of stakeholders and data users, capacities, processes, policies and infrastructure used to capture and analyse data are referred to as the data ecosystem (European Commission, 2017b; StatCan, 2019). An ecosystem includes four main categories of actors (European Commission, 2017b):

- Data generators – the primary sources generating data, whether actively or passively, such as consumers, customers, enterprises, or citizens
- Data service providers – the actors participating in data value creation chain; those collecting, organizing, storing, processing, retrieving, sharing, using, reusing, restricting or destructing data, such as NSOs
- Data business users – those parties, like companies and public administrations who use the insights derived from data analytics to improve performance or quality of life
- End customers – the data consumers, customers, or citizens.

45. The information that can be shared in a data ecosystem can be very diverse in nature (statistical-spatial/geographic, structured-unstructured, digital or hard-copy, qualitive, quantitative, images, text, etc.) and should thus not be limited to any specific type of format.

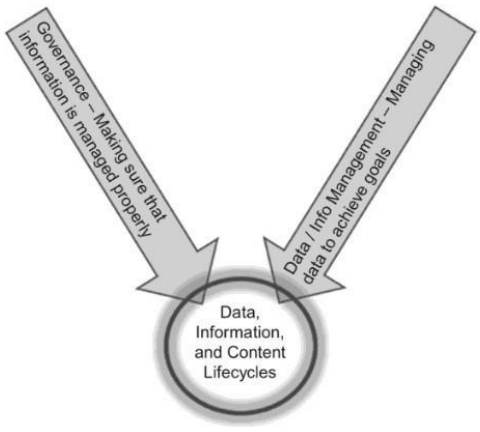
Standards should be open to those interested in establishing the formats and type of information they want to share.

46. The Data Management Body of Knowledge (DMBOK) defines data governance as "the exercise of authority, control, and shared decision-making (planning, monitoring, enforcement) over the management of data assets" (DAMA, 2017). Data governance is about establishing roles and responsibilities – the actual practice of governing. In effect, it is the management of the availability, integrity, interoperability and security of the data stored. It establishes and enforces the policies for access, management, security, sharing and uses of data; identifies the methods and procedures necessary to the stewardship process; and establishes the qualifications of those who would use the data and the conditions under which data access can be granted (Rosenbaum, 2010). This can be in reference to strategic data asset management at the level of the private enterprise or public agency, interdepartmentally, or even at the national level.

47. Data governance can be defined as a system of decision rights and accountabilities for the management of the availability, usability, integrity and security of the data and information, and the resulting regulations, policies and frameworks that provide enforcement. This holistic approach of data governance includes the systems within an enterprise, organisation or government that define who has authority and control over data assets and how those data assets may be used, as well as the people, processes and technologies required to manage and protect data assets (Data Governance Institute, n.d.; IFC, 2021; OECD, 2008, 2019; Plotkin, 2021; StatCan, 2019, 2021b). Data governance enables the coherent implementation and co-ordination of data stewardship activities and increases the capacity (technical or otherwise) to better control the data value chain.

48. Data management is a discipline that directs and supports effective and efficient management of information and data in an organisation, from planning and systems development to disposal or long-term preservation. Data management involves the development, execution, and supervision of plans, practices, concepts, programs, and the accompanying range of systems that contribute to the organization and maintenance of data processes to meet ongoing information lifecycle needs. Data management enables the delivery, control, protection, and enhancement of the value of data and information assets through integrated, user-based approaches. Key components of data lifecycle management include a searchable data inventory, reference and master data management and a quality assessment framework (DAMA, 2017; Government of Canada, 2019; StatCan, 2019, 2020a, 2021b).

Figure 2
Governance 'V' Model *(source: Ladley, 2020)*



49. The governance of data can be conceptualized using the "Governance 'V' Model" (Ladley, 2020), which is useful for the understanding of the subtle differences between data governance and data management. On the left is data governance – the authorities and policies previously discussed. The right side shows the "hands-on" management of data and information (necessary for sound data stewardship). This model emphasizes that data

governance is NOT just a function performed by those who manage data (Ladley, 2020). It also serves to explain the complex relationships between data governance, data stewardship and data management.

50. The "best practices" and principles for data governance and data stewardship include the following understandings: data are a strategic asset; data requires stewardship and accountability; data quality preserves and enhances the value of data; data must be secure and follow privacy regulations; metadata must be standardised, easy to find and of high quality. These concepts have a somewhat cyclical, mutually enabling relationship, but the distinction between them is important, both conceptually and practically.

51. A steward is someone who manages or looks after something on behalf of someone else. In literature the concept of data stewardship is sometimes used to differentiate it from data ownership (McGilvray, 2021). In this report, the term data stewardship emphasizes that public sector data should not be treated as if they are owned by agencies, but that the data should be managed on behalf of and for the benefit of the whole society.

52. For this context, the Task Force proposes the following definition for data stewardship:

> *"Data stewardship means ensuring the ethical and responsible creation, collection, management, use and reuse of data so that they are used for public good and benefit the full community of data users."*

53. Data stewardship:

- Is an approach to data governance that formalizes accountability for managing information resources on behalf of others
- Implements the policies, standards and principles outlined by data governance
- Is enabled through good data governance and data management, which provide the accountability and maintenance or oversight of data assets throughout their lifecycle to ensure their proper care, high quality, data security and confidentiality
- Influences proactive and responsible data practices to help deliver data strategies, maintain trust and promote accountability
- Is expressed through long-term and inter-generational curation of data assets
- Works to support the growing maturity of data policy
- Is made visible through a range of internal and external functions associated with stewardship roles – including data access, security, data quality and standards
- It deals with methods and mechanisms of acquisition, storage, protection, aggregation, deidentification, and procedures for data release, use and re-use, to ensure that the data assets are of high quality, easily accessible and used appropriately
- Is applicable at all scales, from the national or data system level, to the organisation or enterprise level, to the individual or dataset
- Manages and coordinates the interactions of different actors in the system
- Has two main scopes, intra-agency and system-wide data stewardship.

54. While there are similarities in the scope of data governance and stewardship, they are not synonymous. Rather, in practice they seem to occupy two sides of the same coin. The European Union's *Data Governance Act* demonstrates this with its focus on interoperability and sharing, on ensuring public sector data is available for reuse, and on leveraging General Data Protection Regulation (GDPR) to balance the use of personal data with individuals' rights (European Commission, 2016). In enacting and applying the principles of data governance and data stewardship, there can be some overlap. While data governance is the authority and regulatory framework that guides or mandates activities that enable the treatment of data as a strategic asset, data stewardship is the daily enactment of this governance through the implementation of the policies, standards, and principles outlined by data governance (applied data stewardship, see possible roles and responsibilities in Chapter 5). This is why data stewardship often exists as a tangible element of many organisations'

data strategies. Data management is an element or mechanism of data stewardship, and encompasses practical activities like searching for data, sorting data, adding metadata, analysing and cleaning data, etc.

55. As such, both data stewardship and data governance professionals are highly focused on data ethics. “Data ethics” is a newer branch of ethics that focuses on moral problems related to data. Data ethics provides guidelines for the handling of data (from collection to reuse), algorithms and corresponding practices to formulate and support morally positive solutions to data ethics problems. This occurs in full compliance with human rights and works against the risks of misuse and non-use of data for public good. Data ethics refers to the knowledge that allows a person to acquire, use, interpret and share data in an ethical manner including recognizing legal and ethical issues (e.g. security, biases, privacy, confidentiality, and public support or social acceptability) (European Commission, 2020; Floridi and Taddeo, 2016; StatCan, 2020).

56. The data stewardship and governance professionals working in this sphere are subject to data accountability. “Data accountability” is the ongoing liability and responsibility of an individual, enterprise, organisation or public authority regarding the management of data. This principle ensures that appropriate technical and organizational measures and records are in place to establish compliance to data related principles and policy instruments. Examples may include data protection policies, impact assessments, maintaining necessary documentation, data security measures, roles and authorities related to data protection, privacy management frameworks, and appropriate recording and reporting.

IV. Roles and responsibilities for data stewardship

57. The responsibilities of a Data Steward can be linked with a person or an organisation. For example, the National Statistical Office can act as a data steward for the NSS, for a particular aspect of public sector data, or as a national data steward overseeing all public sector data. Within an organisation, there must always be a person or persons (unit) responsible for data stewardship. There can be data stewards, or somebody to whom the stewardship tasks are assigned, in each unit of the organisation dealing with data.

58. Within an agency, the data steward role must have a clear mandate identifying their responsibilities throughout the data life cycle, from collection to dissemination and preservation of the data. Depending on the complexity of the organisation, and the functions associated with the production and processing of data, a specific organizational unit may be designated that includes a position of a Chief Data Steward. It is also vital that this steward be part of a network of data stewards *outside the organization* to promote the use of data as public good for public interests.

59. At the system-wide level, the nature and structure of the government system must be considered. That is, a person should be designated to assume the coordinating role in each sphere of government: one in the executive branch, one in the legislative, one in the judicial and one in each autonomous body. This scheme would be replicated at the state and municipal level according to the applicable models of data governance (such as the centralised, federated and hybrid models discussed in chapter 4 of the full report). It will be useful to also define the role of a data steward in private organizations and companies that own information assets that can be used by others.

60. In the case of NSOs, many positions within their structure are *de facto* data stewards, fulfilling different tasks and responsibilities necessary to implement data stewardship. However, the role of the system-wide data steward must be clearly designated, as this person (or organization) must coordinate with other institutions. Special attention needs to be paid to the relations with privacy protection or personal data protection agencies, to avoid duplication of effort and clarify the division of responsibilities.

61. When using new or alternative public and private data sources from outside the statistical system, the Data Steward in an NSO will need to work closely with other relevant staff, such as persons responsible for data acquisition, data engineering, data quality and metadata. This is explained in more detail in Section 5.2.3 of the full report.

62. Data stewards – at the intra-agency or system-wide level – form a new and essential link in the data value chain. This responsibility can be conceived as three main tasks related to collaboration, ethics and sharing (Verhulst, 2021). Responsible collaboration is necessary to unlock data when there is a public interest case, data must be managed ethically to prevent harm and misuse, and action must be taken to ensure that insight is shared with those who need it, so that it may be translated into meaningful impact.

63. Part of the system-wide stewardship role includes activities NSOs are already undertaking as coordinators of the National Statistical Systems. However, the new role proposed is different in that it dives deeper and covers a wider scope:

- Coordination must be done in a much more complex environment, so further detail and specificity are required using new methods and tools that enable the appropriate scaling: new data sources, including privately held data; the conceptual, methodological and technical challenges of integrating different data sources; increased attention to privacy protection, ethics, inclusivity, public acceptance, etc. New methods and approaches are needed in the whole statistical system and NSOs can and should be a leader in developing and promoting them.
- NSOs' coordination function outside the statistical system is extended, possibly even outside the public sector. This can be done by offering advice and guidance, by providing methods and tools that other agencies can use, or by taking responsibility as data stewards of (some of) the public data holdings.

A. System-wide data stewardship

1. Sound data stewardship – a public sector goal

A. Data strategies and data governance

64. Data stewardship roles and functions assumed by NSOs are often formalised and operationalised in data strategies, describing how NSOs manage data and metadata while providing the best standard of statistical information for the public. The existence of such strategies can be a strong communication asset, arguing the cause of data stewards and explaining to stakeholders how the information is processed, protected and what standards are used. Data strategies are in place to guide NSOs data-related actions and processes, as well as to assure the stakeholders and users that the best information, based on solid analysis and standards, is being provided to users, within a comprehensive framework, with a strong concern about safeguards for data. Data strategies usually encompass such aspects like data collection, data use and re-use, data dissemination and communication, as well as data protection and security. They translate vision and mission into practice by setting strategic objectives of NSOs – which is vital with the data stewardship-driven approach, because those objectives often concern nothing other than the roles and functions of data stewards, i.e., data integration, standard setting, enhancing metadata, spreading solid ethical principles and frameworks within data ecosystems.

65. NSOs have extensive experience with dealing with data throughout its lifecycle. Therefore, the questions of data and information management, data access, privacy protection and data security are high on the respective agendas of governments and public sector departments. Several countries have developed, or are developing, public sector data strategies. Statistical offices have a lot to contribute on these topics and should be involved in this discussion.

Figure 3
Possible coverage of NSO's data stewardship role

66. The data stewardship role that NSOs may take in the public sector can vary as much as the contexts they are situated in. From focusing solely within the NSO, to providing guidance and direction on data management for the whole NSS, or even beyond to the entire public sector. In a context where data stewardship is being applied widely, this can occur within the wider data ecosystem (including towards private data holders); concerning public data; or concerning other producers of official statistics who belong to the national statistical system. This broadest role is also relevant to NSOs in countries where public registers do not exist, since the NSO may still set the direction for government data (e.g. New Zealand). If the focus is relegated to the intra-agency level, data stewardship activities involve a 'core' set of responsibilities for an NSO dealing with sensitive data, and will involve roles that are appropriate for the mandate of an NSO.

67. Figure 3 shows these different responsibilities and their coverage, starting from the outer circle that represents the greatest coverage – data stewardship and management for the whole public sector. Moving inward, each circle describes a successively narrower role, until the inner circle represents the coverage of only one institution - the NSO itself.

68. In cases where the data stewardship role for the whole government sector lies with another institution and the NSO has this role only for official statistics, the NSO should extensively cooperate with the public sector Data Steward. The NSO should look to play an integral part in elaborating standards, quality frameworks and other enabling instruments, given their expertise and the impact these enablers have on official statistics.

69. When data becomes increasingly digitized, the desire to integrate data sets from different organizations for evidence-based decision-making increases. The use of existing data should be managed efficiently, and no data should be collected if there is a public sector organization that has already collected the required data (in accordance with the 'once-only principle'). Thus, the path to better use of data is through increasing the interoperability of data and the use of common technical interfaces, which is facilitated by this public sector-wide alignment.

70. To achieve interoperability, it is necessary to have common guidelines, structures, metadata definitions, quality frameworks, common identifiers etc. in place. The statistical community addressed this problem decades ago by implementing the use of international concepts, classifications and methods to promote the consistency and efficiency of statistical systems (Fundamental Principles of Official Statistics 9) (United Nations, 2014).

71. Increasingly, NSOs are seeking to access privately held data to fulfil their public service mandate for faster and more disaggregated information for government and society. While NSOs are unlikely to be consulted on the interoperability or quality dimensions of such data, the concepts of stewardship described in this paper are even more relevant to the reuse of such data, since the networking, management and legal skills of the data steward will be essential to achieve sustainable access. This role is described in Section 5.2.2 of the full report.

B. The NSO's role in Data Stewardship in the public sector

72. The national context is crucial to the role that the NSO adopts in the data ecosystem. There may already be national bodies with responsibilities for aspects of data governance and management. The NSO will be required to collaborate with these bodies to, at minimum, protect the integrity of the NSO and NSS, and ensure close coordination in the production of official statistics. The NSO may also seek to influence the ecosystem by 'marketing' its skillsets and taking advantage of any roles that are appropriate to an NSO in the broader ecosystem (see also Chapter 7). By nature of the expertise, tools and established partnerships discussed thus far, the NSO could successfully take on (some) responsibilities of a data steward for the whole public sector. However, this may not be feasible in all countries. The following section outlines potential focuses or tasks to advance and manage data stewardship at the public sector level, and some examples of the different approaches taken by countries in this regard. Figure 5 outlines different kinds of data stewardship responsibilities, and examples of the countries pursuing those tasks. Further case studies and examples of initiatives and orientations of various countries follow.

73. It should be noted that the UNECE GAMSO model (2015a), with its focus on strategy and leadership, is ideal for framing the NSO's role in the national context. The GAMSO model defines and outlines the activities that take place within a statistical organization, extending and complementing the Generic Statistical Business Process Model (GSBPM) by adding the activities necessary to support statistical production (2015a).

74. NSOs are uniquely positioned to evaluate the strengths and weaknesses of the national administrative data system. The 'once only' principle is a concept used by the EU that enables efficiency and reduces response burden on citizens, institutions and companies, who are only required to provide certain information to the authorities and administrations *once*. 'Once only' requires that data gaps are identified, and the value of data collected in each silo is maximized through subsequent integration, sharing and analysis. The NSO's role in setting the direction for the system can range from setting up reference classifications to monitoring of the uptake of classifications and standards affecting statistics (e.g. common identifiers). The most evolved or advanced stewardship role would have NSOs' setting the standards for the 'Integrated Data Infrastructure', which are subsequently implemented and monitored across the public data system.

75. Due to their mandates and missions, privacy preservation and social licence are second nature to NSOs. As public sector data quality improves and demands for analysis within individual ministries increase, these discussions are becoming much more relevant to the broader public sector. These public ministries operate in a multidimensional context, where the social impact of their policies have wide-ranging effects in other domains, which contributes to the increasing need for interoperability and data linkage across these domains. While NSOs can meet some of these demands, the NSO may also need to collaborate with other partners while ensuring that overall trust in the system is maintained.

Figure 4.
Summary of system-wide data stewardship work areas, with references to country examples in parentheses.

Data stewardship as a public sector goal

- Contribute to developing and promoting data strategies, policy and principles (Australia, Canada, Estonia, Finland, France, Ireland, Lithuania, Mexico, New Zealand, Switzerland, US)
  - national /public sector data strategies
  - once-only, open data, FAIR, CARE principles
  - Build trust and confidence in the system

Coordination and partnerships

- Coordinate standardisation and harmonisation process, support interoperability (Australia, Mexico, Switzerland)
- Partnerships to improve capabilities, develop new and improve existing products
  - Data sharing arrangements
  - Partnership and licencing agreements
  - Guidance to other agencies (Finland, New Zealand)

Quality assurance and assistance

- Assess data quality in the National Statistical System and promote and provide expertise in data quality to other public data holders (Australia, Estonia, Finland, Germany, Norway, Switzerland)
- Develop and promote quality frameworks (e.g. for administrative registers, privately held data)
- Quality certifications and audit

Data access

- Provide research and information services
  - Access to microdata (Australia, Ecuador, Netherlands)
  - Surveys for other government agencies (Ireland)
  - Information services and analytical support to users (incl. policy makers) (Ireland, Netherlands)
- Build data platforms and dashboards (for data sharing and dissemination) (Australia, Germany, Ireland)
  - Bringing together data from different areas (e.g. Data Centres) (Ireland)
  - Links with geospatial data, geospatial visualization (Croatia)
- Access privately held data sources (Australia)

Methods, tools and capabilities

- Standards and classifications (Australia, Estonia, Ireland, Mexico, Netherlands, New Zealand, Switzerland)
- Infrastructure (e.g. for data integration, access to individual /micro data) (Australia)
  - Data lake (Lithuania)
  - (Meta)data catalogue (Croatia, Switzerland)
  - Common identifiers (Denmark, Ireland, Sweden)
- Data integration
- Data protection
- Statistical methods, data science and machine learning methods (Australia, Netherlands)
- Common approach to data handling (Australia)
- Helping other agencies to build their skills, processes, tools and services (Ireland)
- Improving data literacy (Australia, Estonia)

2. Support to other public sector data providers

A. Coordination and partnerships

76. In many cases, data providers are also data users, since public sector agencies, such as government ministries or central banks usually have policy functions as well as operational ones. At a minimum, information on relevant outputs should be viewed as part of the overall relationship with such agencies. A more proactive approach would involve the analysis of these agencies in the liaison groups governing data flows, where they could have an input into statistical work based on the agencies' data sources. Support for analysts in the public sector can also be organized through networks, which the NSO is ideally placed to facilitate.

B. Data access

77. Many NSOs are considering developments to data and information services as they decide whether to become 'data and statistics' offices. Differential privacy is key to these services. NSOs provide open data and through strategic partnerships, can also facilitate other public bodies to meet their obligations for open data. Researcher access is one of the key areas for development, including the new role many NSOs have taken on of late, in facilitating access to data for COVID researchers. It is also a focus in EU legislation (see Appendix 2), as the Data Governance Act envisages pseudonymised public sector data access for researchers, which may be considered in the stewardship activities of NSO's. Between open and pseudonymised data services, many NSOs also provide anonymised data services to encourage data skills among inexperienced researchers.

C. Quality Assurance (QA) assistance – Methods, tools and capabilities

78. To lead and govern the quality of the data of public organisations, the data steward needs to have several different kinds of abilities and capabilities. Coordination, communication and networking skills are all needed when creating and maintaining national data quality solutions. This is long term work and requires dedicated human resources – resources that are familiar with methodologies as well as those with deep knowledge on the data in question.

79. It is also important to note that the level of data quality required depends on how the data is used. There might also be different kinds of data usages for the same data, leading to different quality assumptions and needs concerning that data. These must be discussed in

detail and with the appropriate subject matter authorities. Specific topics the NSO could consider as part of the stewardship role include:

(a) Metadata. This is crucial for evaluating data quality and needs to be maintained for administrative data within NSOs, as the data were not originally collected for statistical purposes and each source will have its own dimensions of quality depending on mandatory/legislative fields, coverage, etc. The NSO may also consider making the metadata publicly available to researchers, or even more broadly. As a further step, the NSO may define common metadata standards, or collaborate with the body developing the standards to create cohesion and consistency.

(b) Development and promotion of data standards in the community of data owners, such as input and output classifications. As these standards are needed for internal purposes, they may be useful to other bodies collecting similar data to help with use and linkage.

(c) Promotion of unique identifiers as a special category of standards. This is particularly relevant to countries without registers. If coverage of such identifiers is poor, the NSO can promote standards for identifiable data to improve the subsequent level of probabilistic matching.

(d) While the primary concern of NSOs in relation to privately held data to date has been access, quality frameworks for these data also need to be considered. The quality model that applies to public sector data may not be relevant to private sector data. 'Volume, velocity and variety', the defining characteristics of Big Data, will result in different but very informative, statistical products to those derived from traditional administrative sources with high and predictable population coverage levels.

B. Roles associated with intra-agency data stewardship - the role of data steward within an NSO

80. The role of the Data Steward, as it is outlined here, describes the *minimum* activities and responsibilities for the role as it exists in an organization that acquires and processes sensitive data. The NSO context is further explained in upcoming sections of the report by relating the role of the Data Steward to other roles in the organization, based on UNECE GSBPM and RASCI (Responsible, Accountable, Supporting, Consulted, and Informed) matrices. What constitutes the 'maximum list' of activities and responsibilities will depend on the degree of external engagement of the NSO.

81. At minimum, an NSO's data steward would be responsible for:

- Data description (e.g. data catalogue, data dictionary)
- Metadata quality (the responsibility of the data lies with the data holder)
- Data life cycle management
- Data ethics (the Data Steward is the connection between the Data Management officer/Data Protection Officer and the Ethics Committee)
- Data security, protection and confidentiality
- Data audits (monitoring the use of data).

82. The responsibilities of a data steward within an NSO may be categorized by networking activities (both internal and external), technical responsibilities related to data management, and ethical and legal responsibilities.