



Economic Commission for Europe**Conference of European Statisticians****Seventieth plenary session**

Geneva, 20-22 June 2022

Item 8 (d) of the provisional agenda

**Topics of common interest to statistical and geospatial communities -
joint session with the Regional Committee of United Nations Global
Geospatial Information Management for Europe (UN-GGIM: Europe)
Developments in data stewardship****Governance framework in support of data quality – a central
banking perspective****Prepared by Bank of International Settlements***Summary*

The document presents the various data governance projects undertaken by the central bank statistical community. It draws on the various presentations made on the occasion of the High-Level Meeting on Data Governance co-organised by the Irving Fisher Committee on Central Bank Statistics (IFC) and the International Statistical Institute (ISI) in 2019 and hosted by the National Institute of Statistics of Tunisia with the support of the African Union (AU) Commission. The proceedings were published in the IFC Bulletin series (IFC (2021)).

The paper is presented to the Conference of European Statisticians' session on "Developments in data stewardship" for discussion.



I. Introduction

1. Public organisations composing national statistical systems (NSSs), especially central banks' statistical departments, are increasingly aware of the need to follow **strong data governance standards when collecting, managing, disseminating and using official statistics**. Governance frameworks should be holistic, i.e. cover entire organisations including the related principles, policies/procedures, structures, roles and responsibilities, and be an integral part of their strategic plans.
2. Experience shows that well-defined data governance frameworks can be instrumental in supporting central bank statisticians in their work of collecting and analysing data of the highest quality possible and in particular in **ensuring the respect of the Fundamental Principles of Official Statistics** (UN (2013)). Another lesson is the need to complement any institutional level approach to data governance by a broader focus covering the entire production and use of national statistics, including alternative sources. While official statistical institutions may be facing a relative decline in their traditional function of “data collectors”, they could play a greater role as reference custodians for the quality of data used by society. Establishing sound data governance frameworks would effectively support such a new “data curator approach” (Križman and Tissot (2022)).
3. This paper draws on the various data governance projects undertaken by the central bank statistical community. Section 1 shows as how such frameworks can facilitate the working with the (rapidly increasing) amount of data faced by statisticians. Section 2 reviews some key aspects to be considered when actually implementing a data governance framework. Section 3 highlights the value of cooperation in this context. The last part of the paper argues that a data governance framework can bring three important benefits, namely the securing of the role of official statistics as reference information (Section 4), their effective communication (Section 5), and their use for policy purposes (Section 6).

II. Data governance as a way to structure data work

4. Like their counterparts in the NSS, **central bank statisticians are facing an avalanche of data that can be difficult to manage properly**. On the supply side, rapid IT innovation means that databases are becoming larger, easier to access (e.g. “open data”) and also more numerous and heterogeneous – cf the expansion of non-structured data. This calls for more capacities to store, preserve, link, transform and analyse the data. It also requires setting up adequate and standardised procedures for combining all the information together. On the demand side, the potential usefulness of the new information has gone up especially for policymaking. For instance, real time statistics can help understand complex changes to the economy faster and improve how decisions can be made. Moreover, data is an asset that must be adequately safeguarded and controlled, to ensure that it remains safe, reliable and appropriately used, and to protect confidentiality, privacy and intellectual property.
5. These evolutions are clearly **putting pressure on NSSs' IT and budget resources as well as on staff skills**. The implication of the information revolution is not just that more indicators have to be produced by the NSS: for the data to be useful, they need to be transformed and exploited. Moreover, there are specific challenges posed by the integration of alternative sources. These include issues related to data documentation (or “metadata”, to have reassurance on the quality and fitness for purpose of these data), calibration (to ensure that the new indicators compiled provide an accurate view of the underlying reality) and curation (to ensure the adequate organisation and integration of data collected from various sources, including the ones collected by other data communities). In addition, the processing and analytical tools at disposal are subject to rapid innovation, while new data usages are constantly emerging. Furthermore, official statistical bodies have to deal with the specific responsibilities entailed in accessing new, potentially sensitive data.
6. **A data governance framework can alleviate these challenges as it can help to better understand these various issues and address them effectively**. In Portugal, for instance, the central bank has launched an integrated data management programme to make better use of the information available in the whole organisation and rationalise the processes

involved in its collection (IFC (2021)). Concrete objectives are to combine data sets, promote data sharing, allow for regular data exploration and (irregular) experimental analysis, and maximise the usefulness of the data for internal and non-internal (research community) users. The implementation of this programme had two main aspects. First, the definition of a governance model with clear differentiation of responsibilities between the types of staff involved in data management. Second, the development of a logical data architecture encompassing the data, storage facilities, and data exploration tools. Various central banks, including in Thailand and Tunisia, have developed similar approaches (IFC (2021)).

III. Implementing a data governance framework

7. **A key feature of a sustainable data governance framework should be comprehensiveness in covering the entire institution** so that it can become more data-driven. The objective is to make the best use of the data available, enable their access, sharing and integration, and increase overall efficiency and accountability (OECD (2019)). In the BIS case, the framework being developed includes a wide range of data policies, standards and guidelines that can support extracting value from all of the data assets in the institution. Its setup comprises the design of a general data processing architecture (with the structural role played by the Statistical Data and Metadata Exchange (SDMX) information model), the clarification of the business cases across various units, the development of appropriate IT solutions, and the integrated implementation of related software and hardware components. Comprehensiveness has been ensured by defining high level data governance principles; selected key IT underpinnings; and dedicated staff resources (IFC (2021)).

8. Because of their broad scope, **data governance frameworks are often considered as an essential part of the organisations' strategic plans**. Indeed, the framework developed by the BIS is a key building block supporting its overall Innovation 2025 strategy, ensuring its robustness and comprehensiveness. It can also speed up its secure implementation by acting as a potential “car break”, allowing the pursuit of multiple ambitious objectives in parallel – e.g. improved user services, novel data-driven capabilities, (operational and cyber) resilience and business agility.

9. The **European Commission** has also developed a fully-fledged “European strategy for data”, based on four pillars (European Commission (2020)). First, a cross-sectoral governance framework for data access and use, based on a set of high-level guiding principles. Second, key enablers, e.g. investments in data and strengthening Europe’s capabilities and infrastructures for hosting, processing and using data, interoperability. Third, the competences needed (by empowering individuals, investing in skills and in SMEs). And fourth, the development of common European data spaces in strategic sectors and domains of public interest.

10. Turning to the **IMF**, its overarching strategy on data and statistics focusses on three “I”s (IMF (2018)): Integration, to connect the various data provision work streams; Innovation, to take advantage of big data; and Intelligence, to leverage artificial intelligence (AI) to analyse data and statistics. Moreover, two governing bodies have been created: the Standing Committee on Data and Statistics to shepherd the implementation of the data strategy; and the Data Governance Group that develops and monitors the implementation of good management policies, practices and guidelines.

11. **The various experiences above show that success will often depend on a few “best practices”**, including high-level commitment in the organisation, a transparent inventory of all data assets, the design of controls to ensure that agreed procedures are followed and staff resources specialised in data management tasks. In particular, there is a need to differentiate clearly between responsibilities, for instance with the following split: system owners (in charge of using specific information systems); data owners (responsible for information sets in the data warehouse); data custodians (responsible for the safe custody, transport, storage of the data); data stewards (to ensure the quality of the data assets stored, in line with the governance processes established by the organisation; as “day-to-day” contact points between business areas and the information management programme, they often play a key

role in the success of data governance frameworks); data experts (specialised in specific business contents); and data users.

12. It is also important that the **various stakeholders are closely associated, both within the organisation and outside** (other data producers, business communities, cybersecurity experts, etc). Yet the major change is often cultural, by promoting organisational changes and developing effective partnerships between business and technology areas. In other words, data governance is not exclusively a technical issue confined to IT departments, as it applies to all of the organisational structures of an institution. This calls in particular for establishing strong data coordination groups.

IV. The value of cooperation

13. In practice, the establishment of a fully-fledged data governance framework may be hindered by the important resources to be mobilised – an issue of clear importance in view of the budget constraints faced by many NSSs. **One way to go is cooperation.** Setting up formal associations of statisticians as well as more informal forums can be a good first step to share experiences and spread innovative practices. This can be particularly useful in developing regions like Africa, where statistical capacity is limited: (big) data skills can be scarce, and local talents may be attracted by more lucrative positions in other countries. Addressing these challenges requires considering a wide range of issues, i.e. not just training for internal staff, but also the available infrastructure in terms of budget, IT equipment and project management tools, as well as development opportunities (both from an academic and managerial perspectives).

14. **Cooperation can take place at various levels to support capacity building.** First and foremost, it can be organised within the NSS, with a leading role to be played by the NSO and the central bank as key compilers of official statistics; this can be facilitated by establishing proper legal frameworks, effective statistical planning and strong coordination mechanisms. Second, cross-country cooperation can help to publicise “best practices” or “success stories”. For instance, the NSI of Rwanda has decided to set up a Data Science Campus to investigate the use of alternative sources and promote new generation data technologies, leveraging the experience of the UK Office of National Statistics (IFC (2021)). Third, NSS capacity challenges can be addressed through the support of international organisations. As an example, the Partnership in Statistics for Development in the 21st Century, or PARIS21, has developed a comprehensive approach to implement and monitor capacity development to strengthen national statistical governance. Planning has already significantly improved with the development of National Strategies for the Development of Statistics, though limited funding remains a challenge.

V. First benefit: official statistics as reference information

15. **A comprehensive governance framework can help to strengthen the reference role of official statistics.** A key aspect is that they should provide an objective basis for answering questions posed by society, underlining the importance of adequate dissemination to the public. This means presenting evidence-based facts effectively so that they can be understood, accepted and used without question. Data compilers should not only communicate in a clear and understandable way but also present a comprehensive picture of the analytics provided, including the degree of uncertainty associated with the data and techniques used. It also puts a premium on developing statistical literacy in the population, and on ensuring that the public understands what official statisticians are doing and is comfortable with that. These issues have clearly gained prominence with the emergence of big data and “black box evidence”, reflecting the difficulties of communicating on the basis of complex and automatised analytical tools. Lastly, it may require NSS bodies to challenge misrepresentations or bad use of numbers put forward by data producers in the public debate.

16. There is evidence that the role of the NSS in **providing reference information can be instrumental in supporting economic development.** Indeed, the conduct of IMF multilateral surveillance exercises requires cross-country comparable data, and past financial

crises have underlined the importance of providing better and transparent information to markets. The importance of reference statistics is also one key driving force of the push made in recent years to improving countries' statistical production in line with the IMF data dissemination standards.

VI. Second benefit: effective communication of statistics

17. **Governance frameworks can also effectively support the communication of statistics to their users.** For instance, the German NSO has developed a communication strategy to ensure that the data produced are understandable and explicable as well as relevant and easily accessible. This strategy relies on key data governance elements, such as: strong official statistics “brand” especially in terms of quality and trust; improved data access; focus on meeting users' needs, by moving away from a culture of users retrieving information to delivering data to targeted groups; dialogue with external stakeholders: users' forum, newsrooms, social media, feedback surveys; and improved data comprehensibility.

18. **Another successful experience is the open-source approach** adopted by the Statistics Department of the African Development Group as part of its capacity-building program. The increasing demand for official indicators is addressed by a unified statistical portal (the “Africa Information Highway”) based on live open data platforms electronically linking all African countries. This approach has been effective to support the monitoring of the Sustainable Development Goals (SDGs) in Africa. As part of its new strategy on data and statistics, the IMF is also working to establish “global data commons”, i.e. an integrated cloud-based network of country websites publishing data for surveillance.

19. **Yet, in practice communicating statistics also requires paying concrete attention to the specific groups of users involved:** the data compiled are of little use for them if they do not meet their expectations. NSSs' bodies should thus clearly identify and understand users' information needs before deciding on new dissemination projects as well as on respective priorities. In turn, the receivers of statistical information should be able to digest it properly. This calls for helping them to understand, access and make use of the data through adequate visualisation and dissemination tools. To this end, purely quantitative information needs to be transformed to be knowledge accessible for non-statisticians, by providing visualisations, narratives or “stories”, and short “key facts” pages.

20. **Attention has focussed on specific groups of users of statistics, in particular academic researchers.** To make further progress in this area, a number of central banks, NSOs and international organisations have joined forces in the context of the International Network for Exchanging Experience on Statistical Handling of Granular Data (INEXDA) (IFC (2019)). The network's ultimate aim is to help others use granular data for analytical, policy and research purposes. Yet one difficulty in practice is to balance the utility of making information more accessible and the risks involved in terms of confidentiality. A successful example is the Research Data Center and Innovation Lab unit set up by the Bank of Italy to promote the dissemination of microdata and address these challenges (IFC (2021)). Since all the data cannot easily be anonymised, different processes have been developed depending on the type of information considered. For instance, by producing so-called Public Use Files (PUFs) consisting of micro-level records prepared in such a way that individual entities cannot be identified; this allows working on micro data sets but at the price of a loss in information value. Other solutions include the set-up of remote restricted access facilities or the provision of on-site access for selected researchers.¹

VII. Third benefit: increasing data use for policy

21. **A specific aspect of data dissemination relates to policy makers, who have proved to be increasingly important consumers of official statistics.** Data governance frameworks

¹ For the examples related to Eurostat's dissemination of microdata, see ec.europa.eu/eurostat/cros/content/public-use-files-eurostat-microdata-0_en; for the various possibilities set up at Statistics Canada to access micro data, see www.statcan.gc.ca/eng/microdata.

should facilitate the provision of good quality, trustworthy and relevant evidence, and ensure that this evidence is correctly used at the right time. This in turn leads to better decisions by supporting the design, calibration, assessment and modifications to public actions. These tasks can be facilitated by the ongoing data revolution, with the increased availability of new data sources and statistical tools, including predictive analytics such as propensity score matching techniques that can support the assessment of policy effectiveness. Yet three fundamental issues arise. First, what are the key elements to focus on when supporting policy makers in making use of data? Second, what are the implications for the NSS in interacting with public authorities? And third, how can countries learn from each other, since most public policies are undertaken at the domestic level?

22. As regards the first issue, **the starting point is to provide information that is relevant to policymakers.** An important requirement is to have real-time data as much as possible. Timeliness can be enhanced through better coordination among the various stakeholders involved, such as the official data community, media hubs, etc. It can also be improved by streamlining data management processes, for instance with a greater use of alternative, high-frequency indicators as a complement to conventional statistics. Another point is to translate policy requests into information needs, e.g. in terms of types of data and potential sources. It means that users should not be asked which data they want, but rather what their analytical needs are. This calls for a close interaction between official statisticians and public authorities, while still preserving independence that is key for ensuring public trust.

23. A **second issue is how NSS organisations should manage their communications with policy-makers.** Obviously, official statisticians are best placed for selecting and transforming raw data to produce high quality information. But they should also put themselves in the position of policy users and focus on “what are the data sources that can answer that question?” instead of saying “here’s a survey, what can it tell you about the world?” – to quote the former UK NSO Head (CSQ (2019)). Moreover, data experts should not just produce statistics that reflect the reality, but also present these indicators in an easy and attractive way. To this end, NSSs are developing in-house capacity, for instance by setting up dedicated communication units, investing in writing analytical reports and preparing simple statistical outputs that can be easier for use. A last communication aspect is to ensure that authorities are also well equipped to understand the data at stake. Indeed, one important paradox is that despite the increased push for conducting data-driven policies, many decisions continue to be taken without due consideration of statistical basics.

24. The last issue is **how to promote the exchange of experience among countries** since communication with policy makers is mostly organised at the country level and often reflects national idiosyncrasies. One suggestion is to provide more international platforms to provide successful examples of best practices that may not be widely known. For instance, by promoting the dialogue between users and producers of statistics. This can be done by organising “wiki-type based” sharing exercises showcasing successes, setting up adequate user surveys, and outlining the public value of official statistics. Another initiative has been to communicate about statistical “champions”. For instance, the Global Partnership for Sustainable Development Data has engaged a number of such champions to support the SDG data initiative and provides a platform for them to report on progress. One of these champions is Ghana, where the NSI is working across ministries, departments and agencies to improve the way administrative data is collected and connected to other sources so that disaggregated statistics can be derived for SDG monitoring.

VIII. Acknowledgements

25. This paper was prepared by Irena Križman and Bruno Tissot Respectively, former Vice President of the International Statistical Institute (ISI) and Director General of the National Statistical Office (NSO) of Slovenia (2003-2013) (iren.krizman@gmail.com); and Head of Statistics & Research Support, Bank for International Settlements (BIS), and Head of the Secretariat of the Irving Fisher Committee on Central Bank Statistics (IFC) (Bruno.Tissot@bis.org). The views expressed here are those of the authors and do not necessarily reflect those of the BIS, the IFC, the ISI or the NSO.

IX. Selected references

- Civil Service Quarterly (CSQ) (2019): “An interview with John Pullinger, former UK National Statistician”, *Blog*, 30 October.
- European Commission (2020): “A European strategy for data”, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*, COM (2020) 66 final, February.
- IFC (2019): “INEXDA – the Granular Data Network”, *IFC Working Paper*, no 18, October.
- (2021): “Issues in Data Governance”, *IFC Bulletin*, no 54, July.
- International Monetary Fund (IMF) (2018): “Overarching strategy on data and statistics at the Fund in the digital age”, *IMF Policy Paper*, March.
- Križman, I and B Tissot (2022): “Data governance frameworks and the integration of alternative data into official statistics”, *Statistical Journal of the IAOS (SJIAOS)*, forthcoming.
- Organisation for Economic Co-operation and Development (OECD) (2019): “The path to becoming a data-driven public sector”, *OECD Digital Government Studies*, OECD Publishing.
- United Nations (UN) (2013): “Fundamental Principles of Official Statistics”, resolution adopted by the Economic and Social Council, E/RES/2013/21, 28 October.
-