



Economic Commission for Europe**Conference of European Statisticians****Sixty-ninth plenary session**

Geneva, 23-25 June 2021

Item 6 of the provisional agenda

How national statistical systems adhere to the core values of official statistics**Promoting official statistics' core values in the new data ecosystem: the case of Mexico****Prepared by Mexico***Summary*

The document presents the experience of the National Institute of Statistics and Geography of Mexico in promoting the core values of official statistics in the new data ecosystem.

The document is presented to the Conference of European Statisticians' session on "How national statistical organisations adhere to the core values of official statistics" for discussion.



I. Introduction

1. The impact of the technological revolution caused by the convergence of information and communication technologies is in the process of being assimilated by all fields of human activity. The increasing availability of data is a direct result of the transformation that has occurred in the ways in which people live, produce, and access public and private services.
2. Business models consume and generate ever more data that, transformed into information, become regular and frequent inputs for decision-making that is applied to the processes of production and delivery of products and services. The act of consuming those same goods and services generates additional data flows that may be useful to providers and other actors, since it is now possible to “extract” more information by mixing it with other information, and understand very diverse expectations, behaviours and needs.
3. The ability to collect, process, analyze and disseminate this information goes beyond governments or national statistical organizations (NSOs). The data ecosystem has also been transformed in recent years; diverse members participate directly or indirectly in this enriched environment seeking information, and generating, processing, curating, sharing, selling, and preserving data. It is increasingly common to know of organizational initiatives to treat data as strategic assets.
4. This fascinating abundance of data can also be used to enrich the production of statistical information using methodologies, technologies and sound scientific knowledge that allow us to understand multiple phenomena to understand how society works.

II. The role of national statistical offices and the new player in the data ecosystem

5. In general, NSOs around the world have been recognized as “the place of data” for decades and in some countries for more than a century. NSOs enjoy prestige and broad social trust for their discipline, scientific strength, and adherence to principles that offer transparency, independence, and quality of information. They also have known how to evolve and incorporate different demands of society to have reliable information on multiple social, economic, and environmental phenomena; thus, they regularly offer information related to these topics, which is used for the most diverse purposes, from the definition of public policies to academic research.
6. However, given the abundance of data production, today as never before NSOs are urged to analyze the changes in their environment and the ecosystem in which they operate, recognizing the availability of new sources and the presence of other actors interested in playing a role within that space previously dominated –perhaps overwhelmingly– by NSOs.
7. The presence of other actors in the ecosystem is something we must get used to, it is important to recognize their existence and capacity to use big data in the generation of new information, but at the same time they have exposed the dilemma of who owns the data? An issue that must be approached carefully to maintain respect for fundamental rights while increasingly pursuing innovation. Taking advantage of their presence may represent the possibility of developing cooperation projects that could make possible to have alternatives that allow generating more granular, timely and of higher quality statistics.
8. Sponsoring the common adherence of all the participants in that ecosystem to the “same” catalogue of principles can strengthen the generation and use of more statistical information; the transparency of the players in this environment can and should be better recognized by users so that they appreciate its characteristics and have the option to accept or discriminate the information produced by them. The NSO could promote the use of such catalogue and raise the culture of using information by society.

III. The need for common principles for data governance

9. The rules and processes according to which the data ecosystem operates today must evolve, establishing governance mechanisms that allow actors to preserve and increase the trust of the audiences they serve, who in many cases have been the providers of the data themselves. Privacy and the right of individuals to decide about their information must be prioritized; at the same time, it is necessary to facilitate the use of data to generate information that seeks the common good. Perhaps the COVID-19 pandemic –within its multiple teachings– have obliged us to quickly incorporate new data sources; a capacity of response that allows agility and offers transparency so that we are all in a position to judge adherence to ethical principles by transforming this data into useful information for decision-making.

10. Since the end of the last century, countries have had discussions about the importance of regulating the circulation of data, establishing different formulas to protect the right to privacy and protect sensitive data. As a result of these discussions, laws and agreements emerged, specialized data protection agencies were established, and actions were implemented to enforce the exercise of these rights. In this regard, we must underline the importance of the initiatives of the international statistical community to promote the establishment of conventions and harmonized legal frameworks that facilitate legal adaptation in each country and favour international coordination, recognizing that this is a phenomenon of global scope that requires the participation of most nations.

11. During this time, technology and science have also progressed allowing more sophisticated mechanisms to extract value from data. Machine learning and artificial intelligence synthesize this promise of value but alert society and the international community about possible risks and possible behaviours that can steer clear of the ethical and informed use of the data that is harvested, reused and processed. The ends are multiple, the geographies and cultures are diverse, the scope of the companies is even global, beyond the capacity of government of any nation.

12. For this reason, we must pay attention to the initiatives that provoke thinking about the convenience of evolving the social contract around data that allows generating greater value, equity, and trust. NSOs are aware of the value found in very diverse data that are stored by public or private agents. Our statistical community increasingly uses administrative records, satellite images and has opened up to the possibility of harvesting information directly from the internet through methods such as web scrapping as well as the use of machine learning in the production processes of statistics and in the assurance of its quality.

13. Given the growing participation of entities that can generate and share data, as well as organizations with the ability to convert these flows into information for different purposes, it is convenient to take a step forward and promote the role of NSOs as agents of change that allow evolution of the ecosystem on the basis of trust; trust based on transparency and quality; trust supported by principles that are easily distinguishable by users so that they are able to attribute their value to them.

14. Throughout the years, NSOs have gained the trust of society by adhering to values oriented by the ethical fulfilment of their mission: specialized organizations, supported by scientific knowledge, quality, independence of criteria and the transparency of their Fundamental Principles.

15. Adherence to the common principles enshrined in international agreements offers the support to dare to explore the characteristics that we must seek in the development of the data ecosystem, assuming data stewardship roles in cases in which the production of information with public value is favoured, which must be shared under conditions of equity.

IV. The experience of the National Institute of Statistics and Geography

16. In the Mexican case, the National Institute of Statistics and Geography (INEGI) results from the integration of various administrative departments that attended to the responsibilities of producing statistics since the 19th century.

17. The journey to consolidate INEGI as a reliable producer of statistical and geographic information required significant investments in human and organizational capital and the gradual development of an adequate legal framework for the exercise of a specialized subject that should be kept out of the political ups and downs of the government.

18. At the same time that it was strengthening its technical and technological capacities, society granted it autonomy; the Mexican Constitution and the specialized law recognize the importance of having a body that is characterized by offering accessible, objective, transparent and independent information. These same principles are contemplated in international agreements that favor the establishment of a set of rules aimed at gaining confidence due to quality, independence of criteria, the crystalline treatment of the data and the methods used in the calculations and processing and the regularity in the production and dissemination of information.

19. INEGI's legal framework mandates the coordination of the National System of Statistical and Geographical Information (SNIEG), Mexico's National Statistical System (NSS). To leverage INEGI's and SNIEG's legal, institutional, and reputational capacities in the consolidation of a new data ecosystem it is indispensable to strengthen data governance and stewardship practices. The question here is: How to make sure INEGI's role and message is recognized and respected by all stakeholders so that official statistics' core values and principles permeate to actors in the data market, regardless of the sector where they belong, without compromising its prestige?

20. Our approach has started from the most basic but, we believe, effective: To gradually blend INEGI's role as a producer of information with the role of coordinating the SNIEG. This helps connecting, seamlessly, the traditional NSO's main technical function to the responsibility of promoting and implementing policies, standards, principles, and values through effective governance and stewardship of an expanding NSS. This strategy takes advantage of the fact that most stakeholders expect NSOs to execute a leadership role in the new data ecosystem.

21. Here is a specific environment in which communicating core values is paramount towards not just a bigger ecosystem but a more useful one to the people. The SNIEG needs to incorporate to its membership the variety of actors actively participating in the data market. They all need to be part of the system and explicitly assume the Fundamental Principles.

V. The role of the National Institute of Statistics and Geography in the geospatial ecosystem of Mexico

22. INEGI encompasses as its mandate both official statistics as well as the national cartography and geospatial information. Recently, an innovative initiative has been undertaken that could have a positive impact on the topics of improving communication of our core values to our users and in good practices related to upholding actions in decisions and management approaches.

23. One should acknowledge that in the last decades cartography and geospatial information have become a relevant means to communicate and analyze statistical data. Space and time have become an inherent component in the public policy, scientific and business realms. Mexico has a long tradition since the decade of the 70's in the development of Geographic Information Systems and in the use of cartograms to communicate data related to educational planning, electoral redistricting, economic development and environmental issues, among others.

24. Inspired by the Ordnance Survey regarding its efforts in creating a Master Map for Great Britain, INEGI has advanced in the conceptual design and digital implementation of the Mapa Maestro Mexico (MMM) that has as its main purpose to create a community of producers and users of cartography and geospatial information by providing technological artifacts to support interoperability processes.

25. Moreover, as part of this project, INEGI will be sharing with the community some of its core services and products such as: the National Road Network, Urban and Rural Cartography and the National Hydrographic Network.

26. Although there are still societal challenges ahead in terms of the strategy in community building, the INEGI team has already solved some of the key technical issues involved in the MMM project, such as: the design and implementation of unique ID, a standard geodetic framework, technical interoperability standards supported by the Open Geospatial Consortium and metadata best practices.

27. The new language of the MMM will allow INEGI to communicate better its core values as well as ease the communication processes among producers and users of statistical data embraced by a geospatial and time framework. The members of MMM will benefit from an infrastructure to improve decisions based by evidence and to support innovative data management practices.

VI. Challenges for the National Institute of Statistics and Geography and the way forward

28. Still, INEGI faces several challenges in promoting and building the data ecosystem based on official statistics core values. To name a few, we need a whole-of-government strategy and agreement to exploit regularly and consistently administrative registers, and a legal and institutional device to facilitate data flows from private sector sources.

29. Some current efforts that help us tackle those challenges are:

- Adaptation and promotion of the Tool for Quality Assessment of Administrative Registers (HECRA), available at INEGI's webpage, by which any public entity may assess the feasibility of using their administrative registers for statistical or geographical purposes, identify gaps to be breached, and implement improvement plans;
- *Nowcasting* of an indicator of economic activity, using timely economic and financial variables, and high frequency alternative sources such as Google Trends and a mobility indicator;
- *Nowcasting* of an indicator of manufacturing activities, using electric energy consumption data provided by the Federal Electricity Commission;
- Data flows from automobile and heavy-duty vehicles, for which INEGI serves as the outlet for those administrative registers;
- Memorandums of understanding with commercial banking institutions that will allow us to link data to traditional statistical and geographical information, and
- Sentiment analysis using Twitter data, which by recollecting daily tweets provides results by time and state.

30. In 2014, INEGI created a high-level committee for quality assurance. The principles and guidelines promoted by this committee are aligned with the UN National Quality Assurance Framework (NQAF), which includes the UN Fundamental Principles of Official Statistics. In 2020, the internal regulations were renewed to align to the new NQAF manual and to guide the committee decisions towards three principal actions: standardize procedures, assess quality principles, and promote quality improvement.

31. The main purpose of the committee is to give certainty to final users regarding data quality (relevance, accuracy, timeliness and punctuality, accessibility, coherence and comparability, metadata), and methodological soundness in a transparent, impartial and objective environment. The committee has approved a set of indicators to assess data quality, some of which area provided to the final users. It also has implemented the GSBPM in all statistical and geographical programs to standardize process reports and main guidelines. Finally, it has developed a system to report design changes and the results of statistical trials. The whole quality strategy has included training courses and internal communication activities.

32. Another example of improving the regulatory framework for the SNIEG is a state-of-the-art statistical confidentiality policy that soon will be approved by INEGI's board, and will be complemented with implementation guidelines and training courses.

VII. Conclusion

33. In the forthcoming years we expect to expand the quality strategy to the SNIEG. We already started to renovate policies to align them with the GSBPM and the new quality principles and guidelines. At INEGI we are convinced that we have to promote the quality principles and provide guidelines for data production among all producers, including all participants in the statistical and geospatial ecosystem.
